Revenue Management Models in the Manufacturing Industry

by

Tieming Liu

B.S., Process Automation and Automatic Instruments (1997)
Tsinghua University
M.S., Control Theory and Control Engineering (2000)
Tsinghua University
M.S., Industrial Engineering and Management Science (2001)
Northwestern University

Submitted to the Department of Civil and Environmental Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in the Field of Transportation

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2005

Author ....
Department of Civil and Environmental Engineering
July 25, 2005

Certified by ....
David Simchi-Levi
Professor of Department of Civil and Environmental Engineering and
Engineering Systems Division
Thesis Supervisor

Accepted by ....
Andrew Whittle
Chairman, Departmental Committee on Graduate Students

Revenue Management Models in the Manufacturing Industry

by

Tieming Liu

Submitted to the Department of Civil and Environmental Engineering
on July 25, 2005, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in the Field of Transportation

Abstract

In recent years, many manufacturing companies have started exploring innovative revenue management technologies in an effort to improve their operations and ultimately their bottom lines. Methods such as differentiating customers based on their sensitivity to price and delays are employed by firms to increase their profits. These developments call for models that have the potential to radically improve supply chain efficiencies in much the same way that revenue management has changed the airline industry.

In this dissertation, we study revenue management models where customers can be separated into different classes depending on their sensitivity to price, lead time, and service. Specifically, we focus on identifying effective models to coordinate production, inventory and admission controls in face of multiple classes of demand and time-varying parameters.

We start with a single-class-customer problem with both backlogged and discretionary sales. Demand may be fulfilled no later than N periods with price discounts if the inventory is not available. If profitable, demand may be rejected even if the inventory is still available. For this problem we analyze the structure of the optimal policy and show that it is characterized by three state-independent control parameters: the produce-up-to level (S), the reserve-up-to level (R), and the backlog-up-to level (B). At the beginning of each period, the manufacturer will produce to bring the inventory level up to S or to the maximum capacity; during the period, s/he will set aside R units of inventory for the next period, and satisfy some customers with the remaining inventory, if expected future profit is higher; otherwise, s/he will satisfy customers with the inventory and backlog up to B units of demands.

Then, we analyze a single-product, two-class-customer model in which demanding (high priority) customers would like to receive products immediately, while other customers are willing to wait in order to pay lower prices. For this model, we provide a heuristic policy characterized by three threshold levels: S, R, B. In this policy, during each period, the manufacturer will set aside R units of inventory for the next period, and satisfy some high priority customers with the remaining inventory, if expected future profit is higher; otherwise, s/he will satisfy as many of the high priority customers as possible and backlog up to B units of lower priority customers.

Finally, we examine production, rationing, and admission control policies in manufacturing systems with both make-to-stock(MTS) and make-to-order(MTO) products. Two models are analyzed. In the first model, which is motivated by the automobile industry, the make-to-stock product has higher priority than the make-to-order product. In the second model, which is motivated by the PC industry, the manufacturer gives higher priority to the make-to-order product over the make-to-stock product. We characterize the optimal production and order admission policies with linear threshold levels. We also extend those results to problems where low-priority backorders can be canceled by the manufacturer, as well as to problems with multiple types of make-to-order products.

Thesis Supervisor: David Simchi-Levi
Title: Professor of Department of Civil and Environmental Engineering and Engineering Systems Division

Acknowledgments

Though many days were spent envisioning this moment, it now surpasses my imaginable expectations, and that is due to the countless people that I owe greatly for their support and guidance throughout this endeavor.

To my advisor, Professor David Simchi-Levi, for being an inspiring and patient mentor as well as a never-failing source of ideas and enthusiasm throughout all stages of my work. Your wisdom and direction have created an exceptional learning environment and made my research so fruitful.

To my thesis committee members, Professors Stephen Graves and Patrick Jaillet, who generously offered their time, provided their insights, and shared their ideas. Your inputs in the thesis are highly appreciated.

I also want to express appreciation to my research collaborators outside of MIT. Professor Seyed Iravani at Northwestern University, Professor Julie Swann at Georgia Institute of Technology, and Dr. Laura Wynter at the IBM Watson Research Center all provided valuable inputs and advices for my research.

I would like to thank my friends at MIT, for their generous help in my life and study: Xin Chen, Shan Lan, Chunmei Liu, Hao Wang, Jin Yi, ..., just to name a few. I would also like to thank my officemates, Song Gao, Hai Jiang, and Michael Hanowsky for proving an enjoyable working environment.

Finally, I am indebted to my family for their support. My parents' selfless love and support never failed and is the foundation for all I have done and the inspiration for all I might achieve. My wife Yu Mao gives me endless love and encouragement. What I have accomplished at MIT would not be possible without her encouragement and support.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivations

In the past three decades, revenue management - integrating price, inventory control, and quality of service - has been applied to more and more industries. Techniques such as market segmentation have been successfully applied in airlines, hotels and car rental agencies [25]. For example, in the airline industry, companies differentiate business travel customers from leisure travel customers by setting advance purchase and Saturday night stay requirements, and they offer different "fare products" for the same travel in the same O-D market. In the retail industry, to name another example, dynamic pricing techniques can provide significant improvements in profitability [14].

In recent years, scores of manufacturing companies have started exploring innovative revenue management technologies in an effort to improve their operations and ultimately their bottom lines. Methods such as differentiating customers based on their sensitivity to price and delays are employed by firms to increase their profits. These developments call for models that have the potential to radically improve supply chain efficiencies in much the same way that revenue management has changed the airline industry. In other words, we need to extend the research area of revenue management from traditional industries with perishable products, such as airlines and hotels, to the manufacturing industry with non-perishable products.

For instance, no company underscores the impact of customized pricing strategies more than Dell Computers. Dell separates its customers into different classes, such as home users, small business, large business, government, education, and so on, and for the same products, Dell may charge different classes of customers different prices. A more careful review of Dell's strategy, see [1], suggests that even the price of the same product for the same industry is not fixed; it may change significantly over time.

Dell is not alone in its use of sophisticated market segmentation and pricing strategies. Other companies, such as IBM, HP, and Gateway, are also applying the same practice. Ford Motor Co., to name another example from a different industry, credits $3 billion in growth between 1995 and 1999 for the effort of using pricing strategies to match supply and demand and target particular customer segments [21].

However, research on the application of revenue management in the manufacturing industry is still in its early stages. There are a number of characteristics that distinguish general manufacturing industries from the industries mentioned previously, including the non-perishability of products and the ability to vary production levels. More importantly, manufacturing differs from most retail environments in its reordering and capacity characteristics. Specifically, manufacturers typically have limited production capacity while retailing often involves a single large order at the start of a selling season.

These developments call for models to effectively coordinate production, inventory and admission controls in the face of time-varying parameters, such as dynamically changing prices, stochastic customer demands, as well as fluctuating production capacities and production costs. This is precisely the focus of Chapters 2 and 3 in this thesis.

Another important trend in the manufacturing industry is customized production, the make-to-order (MTO) environment. Customized production not only gives more satisfaction to customers, but it also helps manufacturers eliminate finished goods inventory. However, make-to-order environments suggest important challenges associated with matching fixed production capacity with highly variable demand. Specifically, an MTO system implies periods where the facility is idle and other times in which a large number of orders are waiting for production.

Of course, while some customers enjoy customizing their products, other customers would accept a standard configuration in order to receive their products immediately. To satisfy these customers, manufacturing companies also produce a variety of standard products to stock, the make-to-stock (MTS) environment. For example, while selling customized PCs through the internet, Dell also frequently provides promotions for some low-end products to attract more customers, and for these promotional products, customers usually have very little flexibility in product configurations. This gives raise to a combination of a make-to-stock/make-to-order (MTS/MTO) environment that allows manufacturer to better manage their production capacity and increase expected profit.

The application of make-to-stock/make-to-order manufacturing systems is also important among part suppliers who face demands from both original equipment manufacturers (OEMs) and the aftermarket. For example, in the automobile industry, a part supplier sells its products to automotive assembly plants for installation into new vehicles, as well as repair shops for replacement in old vehicles [5]. OEM and aftermarket demands are both important to the part supplier. OEM demands guarantee high utilization of the production capacity, while aftermarket demands bring high profit margins to the supplier. OEM sales are based on long-term contracts, and they are produced under the make-to-stock mode. In contrast, aftermarket items are produced under the make-to-order mode due to their large variety.

These developments call for models that integrate production, sequencing and admission decisions for a hybrid production system with both make-to-stock products and make-to-order products. This is exactly the focus of Chapter 4 in this dissertation.

1.2 Literature Review

One stream of literature related to our research is inventory control problems with production capacity constraints. Federgruen and Zipkin [12][13] studied the production and inventory control strategies for a single-class-customer problem with production capacity constraints and backlogged sales, and they showed that a modified order-up-to policy is optimal.

Chan, Simchi-Levi and Swann [6] examined a single-class-customer problem with production capacity constraints and discretionary sales. In discretionary sales, inventory may be set aside in the current period to satisfy high-price demand in the future. The authors analyze the structure of the optimal policy and show that it is characterized by two parameters: the order-up-to level and the reserve-up-to level. The model presented in Chapter 2 is an extension of this line of research.

Another stream of related literature considers multiple classes of customers with stochastic demand. Specifically, single-product models with multi-class customers in a make-to-stock manufacturing systems, the so-called stock rationing problems, have been studied in various contexts since the late 1960's. Topkis [29] studied a multi-class-customer problem with lost sales under the periodic review model, and proved that the optimal policy has a threshold structure in which some level of stock is reserved for future, more valuable customers. Nahmias and Demmy [24] studied a two-class-customer problem with backlogged sales and analyzed the cost savings due to stock rationing under both the periodic review model and the continuous review model. Cohen, Kleindorfer and Lee [7] studied a two-class-customer problem with lost sales under a periodic review model. They considered an (s, S) type policy and developed a heuristic algorithm to minimize the expected cost. Melchiors, Dekker and Kleijn [23] examined a two-class-customer problem with lost sales under a continuous review model. They develop an optimal critical level policy in the context of an (r, Q) inventory model for Poisson demand and deterministic production lead time. In these papers, the production capacity constraints were not considered.

Stock rationing problems with production capacity constraints have been considered only in the last few years. Ha [17] studied a multi-class-customer problem with lost sales. He assumed exponential production time and Poisson arrival demands, and he characterized the optimal policy with multi threshold levels. In a related study, Ha [18] also considered a two-class-customer problem with backlogged sales, and he characterized the optimal policy with switching curves. De Vericourt, Karaesmen, and Dallery [8] studied a multi-class-customer problem with backlogged sales, and they characterized the optimal policy using structural multiple threshold levels. Extensive reviews for the stock rationing problems can be found in Kleijn and Dekker [22]. The models presented in Chapter 3 is related with this stream of research.

Finally, limited research exist on make-to-stock/make-to-order environments. On the other hand, there is a significant amount of literature addressing control problems for a make-to-order manufacturing system that produces multiple products. The optimality of simple-index rules, e.g., the $c\mu$ rule, has been extensively studied when set-ups are not required to switch from producing one product to another. Please refer to Baras et al. (1985)[2], Buyukkoc et al. (1985)[4], Varaiya et al. (1985)[30] , Walrand (1988)[32], Gittins (1989)[15] and the references therein. For the problem with set-ups, researchers focused on heuristic policies and performance evaluation. Please see Federgruen and Katalan (1996)[11], Duenyas and Van Oyen (1996)[9], Reiman and Wein (1998) [27] and the references therein.

All the papers above focused on dynamic production and sequencing problems for either make-to-order systems or make-to-stock systems; however, none of them considered a hybrid manufacturing system, and admission control was rarely studied. Carr and Duenyas [5] have been the first to consider both the production and admission decisions for a make-to-stock/make-to-order system, where the make-to-stock orders have higher priorities, and they studied a continuous review model and found an optimal policy characterized by monotonic nonlinear switching curves. The model in Chapter 4 in this dissertation is directly related to the model in Carr and Duenyas.

1.3 Contributions

In this dissertation, we study revenue management models where customers can be separated into different classes depending on their sensitivity to price, lead time, and service. Specifically, we focus on identifying effective models to coordinate production, inventory and admission controls in the face of multiple classes of demands and time-varying parameters, such as dynamically changing prices, stochastic customer demands, as well as fluctuating production capacities and production costs.

We start (Chapter 2) with a single-class-customer problem with both backlogged and discretionary sales. Demand may be fulfilled no later than N periods with price discounts if the inventory is not available. If profitable, demand may be rejected even if the inventory is still available. For this problem we extend the base-stock policy and show that the optimal policy is characterized by three state-independent control parameters: the produce-up-to level (S), the reserve-up-to level (R), and the backlog-up-to level (B). At the beginning of each period, the manufacturer will produce to bring the inventory level up to S or to the maximum capacity; during the period, s/he will set aside R units of inventory for the next period, and satisfy some customers with the remaining inventory, if expected future profit is higher; otherwise, s/he will satisfy customers with the inventory and backlog up to B units of demands.

Next, Chapter 3 studies a single-product, two-class-customer model in which demanding (high priority) customers would like to receive products immediately, while other customers are willing to wait in order to pay lower prices. For this model, we provide a heuristic policy characterized by three threshold levels: S, R, B. In this policy, during each period, the manufacturer will set aside R units of inventory for the next period, and satisfy some high priority customers with the remaining inventory, if expected future profit is higher; otherwise, s/he will satisfy as many of the high priority customers as possible and backlog up to B units of lower priority customers.

Finally, in Chapter 4, we study production, rationing, and admission control policies in manufacturing systems with both make-to-stock(MTS) and make-to-order(MTO) products. Two models are analyzed. In the first model, which is motivated by the automobile industry, the make-to-stock product has higher priority than the make-to-order product. In the second model, which is motivated by the PC industry, the manufacturer gives higher priority to the make-to-order product over the make-to-stock product. We characterize the optimal production and order admission policies with linear threshold levels. We also extend those results to problems where low-priority backorders can be canceled by the manufacturer, as well as to problems with multiple types of make-to-order products.

Chapter 2

Single-Class-Customer Problem with Backlogged and Discretionary Sales

2.1 Introduction

In this chapter, we study the coordination of production and admission decisions in a horizon with multiple planning periods and stochastic demand. In this problem, a vector of prices is given at the beginning of the horizon, and production and admission decisions are made at the beginning of each period.

We consider both backlogged and discretionary sales. It is already widely known that backlogged sales can result in more profit, but discretionary sales are not discussed so often. In discretionary sale, the inventory is set aside to satisfy future demands, even though the decision means losing potential sales in the current period. Discretionary sale may be profitable when it is likely to generate a larger income in the future, which would typically occur if the price in the future is higher.

2.2 Model

We analyzed a model with a single product and a single demand class in multiple planning periods, which are indexed consecutively from $1, \ldots, T$. The manufacturer's production capacity in period t is denoted by q_t and the production cost is c_t per unit, for $t = 1, \ldots, T$. The inventory cost is h_t per unit, charged for inventory carried over from period t to period $t+1$. The salvage value at the end of the horizon is v per unit, and it is less than the selling price in the last period.

We assume that the amount of demand is a non-stationary, time-dependent, general stochastic function, D_t, but we do not assume a particular distribution. Let the cumulative demand distribution for a given period t be Φ_t, and let the corresponding probability density function be ϕ_t. The product's selling price at period t is p_t. We assume that customers are willing to wait for up to N periods if the inventory is not available, and backorders must be fulfilled before new orders are satisfied. We assume that the manufacturer can either backlog or reject demands. If a backorder is carried over period t, the manufacturer incurs a backlogging penalty, b_t. If a demand is rejected in period t, the manufacturer incurs a unit of rejection penalty of r_t.

The manufacturer determines the production quantity at the beginning of each period. Let I_t represent the inventory level at the beginning of period t before production, and let Y_t be the inventory level after production and before the new demands are realized. Therefore, $Y_t - I_t$ is the production quantity. At the same time, the manufacturer also determines the minimum amount of inventory to reserve for future sales, R_t, and the maximum amount of backorders to the next period for fulfillments, B_t.

The sequence of events in each period is as follows: at the beginning of a period, the manufacturer checks the inventory level and decides on the production quantity; products arrive in zero lead time, and then the manufacturer fulfills the backlogged demands with the available inventory according to the FCFS (first come first serve) policy; the manufacturer decides on the minimum amount of inventory R_t to reserve for future sales, and on the maximum amount of cumulative backlogged demands B_t

that will be transferred to the following periods; and as demands arrive during the period, the manufacturer realizes, backlogs, or rejects demands with respect to the FCFS policy.

Let Q_t be the amount of acceptance capacity in period t, let H_t be the amount of holding inventory at the end of period t, and let L_t be the actual amount of backorders transferred from period t to $t+1$.

$$
\begin{aligned}
Q_t &= \begin{cases} B_t + (Y_t - R_t)^+ & if\ Y_t \geq 0 \\ (B_t + Y_t)^+ & if\ Y_t < 0 \end{cases} \\
H_t &= \begin{cases} \max(R_t, Y_t - D_t) & if\ Y_t \geq R \\ Y & if\ 0 \leq Y_t < R \\ 0 & if\ Y_t < 0 \end{cases} \\
L_t &= \begin{cases} \min(B_t, (D_t - (Y_t - R_t)^+)^+) & if\ Y_t \geq 0 \\ \min(B_t, D_t - Y_t) & if\ Y_t < 0 \end{cases}
\end{aligned}
$$

Considering the constraints, $0 \leq R_t \leq Y_t^+$, $Y_t^- \leq B_t \leq R_t + q_{t+1} + \ldots + q_{t+N}$, the above functions can be simplified.

$$
\begin{aligned}
Q_t &= B_t + Y_t - R_t \\
H_t &= \max(R_t, Y_t - D_t) \\
L_t &= \min(B_t, (D_t - Y_t + R_t)^+)
\end{aligned}
$$

We use the phrase "profit-to-go" to refer to the expected profit from the current period until the end of the time horizon. Let $J_t(I_t)$ be the profit-to-go at the beginning of period t before production with initial inventory level I_t, and let $G_t(Y_t)$ be the expected profit to go after production with Y_t units of product available. The first and second derivatives of $J_t(I_t)$ are denoted in the following manner, respectively: $J_t'(I_t)$ and $J_t''(I_t)$. Given a vector of price, the profit-to-go functions under the Delayed Production Strategy are:

$$
J_t(I_t) = \max_{Y_t : \max\{0, I_t\} \leq Y_t \leq I_t + q_t} \{-c_t(Y_t - I_t) + G_t(Y_t)\}. \tag{2.1}
$$

The first term in (2.1) is the production cost, and the second term is the expected profit-to-go with Y_t units of products available after production but before satisfying new demands. $G_t(Y_t)$ is defined as:

$$G_t(Y_t, R_t, B_t) = \max_{0 \le R_t \le Y_t^+, Y_t^- \le B_t \le R_t + q_{t+1} + \ldots + q_{t+N}} \int \{p_t \min(D_t, Q_t) - h_t H_t \\ -r_t(D_t - Q_t)^+ - b_t L_t + J_{t+1}(H_t - L_t)\} d\Phi_t \tag{2.2}$$

Finally, we let $J_{T+1}(I_T) = v \cdot I_T$, which is the expected salvage value of leftover inventory.

The first term in (2.2) is the selling revenue. The second term is the inventory holding cost. The third term is the penalty associated with rejected demands. The fourth term is the penalty associated with backlogged demands. The last term represents the profit-to-go from the end of this period forward.

Next, we show that under an optimal policy, in any period, either the amount of reserved inventory equals zero or the amount of backlogged demands equals zero. They cannot both be positive.

Lemma 2.1 *In any optimal policy, we have $R_t \cdot B_t = 0$, for $t = 1, 2, ..., T$.*

Please refer to the appendix for the proof. The intuition behind the lemma is very simple. Suppose that $R_t > 0$, which means that the manufacturer may reject some demands in period t in order to reserve some inventory to the future, then it would not make sense for the manufacturer to use future inventory to fulfill any demand in period t, thus we will have $B_t = 0$. The intuitive explanation of the case with $B_t > 0$ is similar.

With Lemma 2.1, the structure of the optimal policies can be greatly simplified. The manufacturer can choose one out of the two policies: either the reserve-inventory policy ($R_t \ge 0, B_t = 0$) or the backlog-demand policy ($B_t \ge 0, R_t = 0$), thus,

$$G_t(Y_t) = \max\{G_t^R(Y_t), G_t^B(Y_t)\}.$$

$G_t^R(Y_t)$ and $G_t^B(Y_t)$ represent the profit-to-go with Y_t units of products available after production under the optimal reserve-inventory policy or the optimal backlog-demand policy respectively.

$$G_t^R(Y_t) = \max_{R_t: 0 \le R_t \le Y_t} g_t^R(Y_t, R_t) \tag{2.3}$$

$$G_t^B(Y_t) = \max_{B_t: 0 \le B_t \le q_{t+1}} g_t^B(Y_t, B_t) \tag{2.4}$$

$g_t^R(Y_t, R_t)$ indicates the profit-to-go with Y_t units of products available after production and R_t units are reserved for the next period under the reserve-inventory policy; $g_t^B(Y_t, B_t)$ indicates the profit-to-go with Y_t units of products available after production and at most another B_t units of demands can be backlogged under the backlog-demand policy.

$$\begin{aligned} g_t^R(Y_t, R_t) = \; & \int \{p_t \min(D_t, Y_t - R_t) - h_t \max(R_t, Y_t - D_t) \\ & -r_t(D_t - (Y_t - R_t))^+ + J_{t+1}(\max(R_t, Y_t - D_t))\} d\Phi_t \end{aligned} \tag{2.5}$$

$$\begin{aligned} g_t^B(Y_t, B_t) = \; & \int \{p_t \min(D_t, Y_t + B_t) - h_t(Y_t - D_t)^+ - r_t(D_t - Y_t - B_t)^+ \\ & -b_t \min((D_t - Y_t)^+, B_t) + J_{t+1}(\max(-B_t, Y_t - D_t))\} d\Phi_t \end{aligned} \tag{2.6}$$

2.3 Optimal Policy

In this section, we show that functions $g_t^R(Y_t, R_t)$ and $g_t^B(Y_t, B_t)$ are quasi-concave in R_t and B_t respectively, and each of them has a unique unconstrained optimizer that is independent of the inventory level Y_t. The expected profit functions $J_t(I_t)$ and $G_t(Y_t)$ are concave functions of inventory I_t and Y_t respectively. These results are summarized in the following theorem (see the appendix for the proof).

Theorem 2.1 *In any period $t = 1, ..., T$,*

- *$g_t^R(Y_t, R_t)$ is a quasi-concave function of R_t;*
- *$g_t^B(Y_t, B_t)$ is a quasi-concave function of B_t;*
- *$G_t(Y_t)$ is a concave function of Y_t;*
- *$J_t(I_t)$ is a concave function of I_t;*
- *The unconstrained optimizers of $g_t^R(Y_t, R_t)$ and $g_t^B(Y_t, B_t)$, R_t^* and B_t^*, are independent of inventory level Y_t, where*

$$R_t^*(Y_t) = \arg\max_{0 \le R_t}\{g_t^R(Y_t, R_t)\}, B_t^*(Y_t) = \arg\max_{0 \le B_t}\{g_t^B(Y_t, B_t)\}. \tag{2.7}$$

Let R_t^c and B_t^c be the constrained optimizer of $g_t^R(Y_t, R_t)$ and $g_t^B(Y_t, B_t)$, respectively; $R_t^c = \min(R_t^*, (Y_t)^+)$, $B_t^c = \max(Y^-, \min(B_t^*, \sum_{i=1}^N q_{t+i})$. Theorem 2.1 implies the optimal policy, and thus we have the following corollary.

Corollary 2.1 *Given a vector of prices, there exists an optimal policy characterized by an order-up-to level (S_t^*), a reserve-up-to-level (R_t^c) and a backlog-up-to level (B_t^c).*

The optimal policy is characterized by three parameters, and we denote it as the (S, R, B) policy. For the order-up-to level, the policy is to produce enough to bring the inventory level up to S_t^* if there is sufficient capacity, otherwise produce to the maximum capacity. If the reserve-up-to level is positive, the manufacturer will set aside R_t^c units of inventory for the future, s/he will only accept $Y_t - R_t^c$ units of demands; if the backlog-up-to level is positive, the manufacturer will accept up to $Y_t + B_t^c$ units of demands.

2.4 Computational Analysis

In the following, we report a computational study conducted to obtain insights about the benefits of the optimal policy. Our goal is to examine the relative performance of the (S, R, B) policy and identify the situations where the (S, R, B) policy can provide significant increases in profit.

The benchmark we use is a traditional policy, in which the manufacturer uses the modified basestock policy (S policy) and unsatisfied demands are lost. We compare this traditional policy to our (S, R, B) policy using the benchmark ratio. We define the benchmark ratio, *Profit Potential*, as

$$Profit\ Potential = \frac{V_{(S,R,B)}}{V_S} - 1, \tag{2.8}$$

where $V_{(S,R,B)}$ is the expected profit under the optimal policy, and V_S is the expected profit under the traditional basestock policy.

In the computational study, we examine cases with variations on demand seasonality, demand uncertainty, production capacity, and customers' waiting time.

For demand seasonality, we are interested in the variability of the deterministic portion of demand over the entire horizon. We use the coefficient of variation of $E(D_t)$ over the horizon, $CV_S = s_{t=1,\ldots,T}(E(D_t))/Dem^*$, to measure the demand seasonality, where Dem^* denotes the expected average demand over the horizon. Production capacity is constant for a particular problem, while it is allowed to vary from problem to problem from 50% to 100% of the expected average demand over the horizon. Please refer to Table 2.1 in Appendix 2.7 for the data in the case of $CV_S = 0.115$.

For demand uncertainty, we examine the impact of the epsilon component of the demand function. We define the coefficient of variation of demand uncertainty in a given period as $CV_U = s(D_t)/E(D_t)$, where s denotes the standard variation, and E denotes the expected value. For a given test case, the coefficient variation of demand uncertainty is the same in each period.

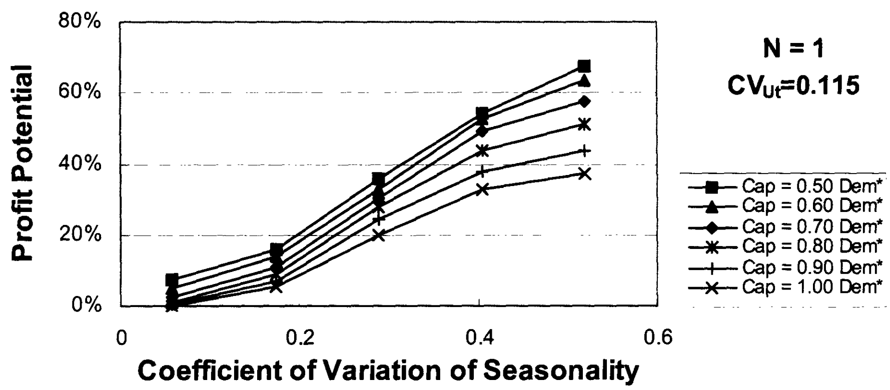


Figure 2-1: Impact of Demand Seasonality

Our computational study reveals a number of valuable insights. The (S, R, B) policy significantly increases profit, particularly when deterministic demand (seasonality) is highly variable and when capacity is tight. The relative performance of the (S, R, B) policy increases as seasonality increases and as capacity decreases, as indicated in Figure 2-1. In the optimal policy, the manufacture can move some inventory between periods to better match the demand. Therefore the optimal policy becomes more effective when demand seasonality is high or when production capacity is tight.

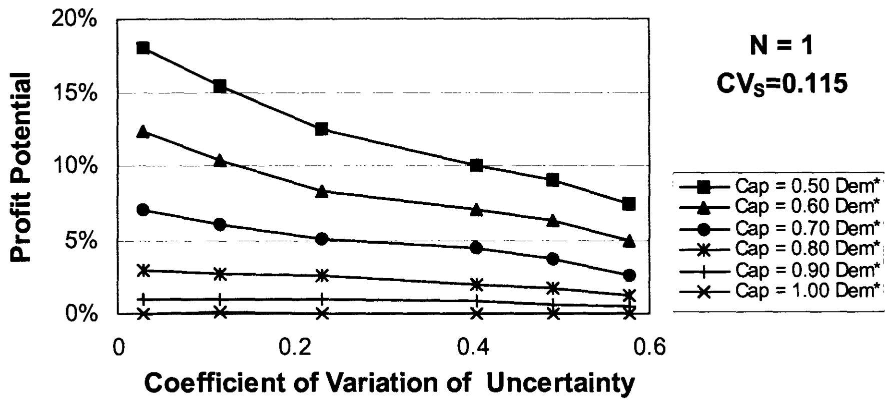


Figure 2-2: Impact of Demand Uncertainty

It is also interesting to find that relative performance of the new policy decrease as demand uncertainty increases (Figure 2-2). One way to explain this is that when demand uncertainty increases, it is less incentive to transfer demands or inventory to the next periods due to the increasing risk. The computational analysis testifies that both the optimal reserve-up-to level and the optimal backlog-up-to level decrease as demand uncertainty increases.

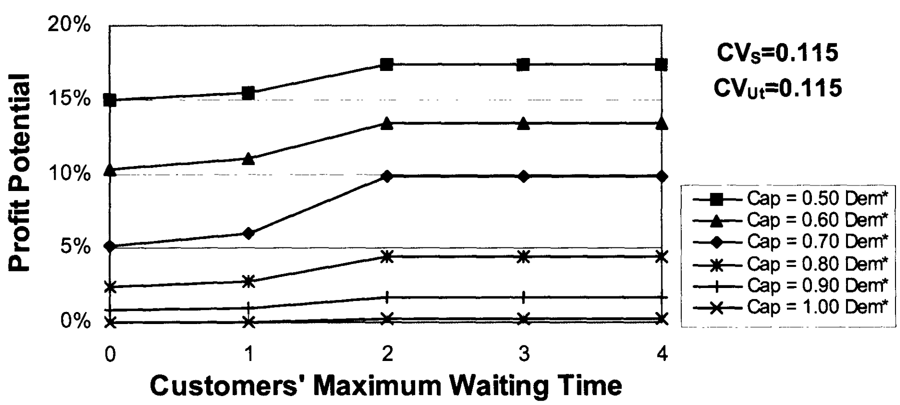


Figure 2-3: Impact of Customers' Maximum Waiting Time

In Figure 2-3, we can see that the relative performance of the (S, R, B) policy increases as customers' maximum waiting time increases from one to two periods, but after that the marginal benefit by increasing customers' waiting time becomes very small. When the production capacity is tight, the profit potential also increases by extending customers' waiting time from two to three periods.

2.5 Concluding Remarks

In this chapter we study the optimal production and admission control policy for a single-class-customer problem in a multi-period horizon. By proving that the profit-to-go functions are concave throughout the planning horizon, we show that a modified order-up-to policy, the (S, R, B) policy, is optimal for this problem. In such a policy, S is the base-stock level, R is the minimum amount of inventory that needs to be reserved for the next period, and B is the maximum amount of demand that will be delayed to the next period. Computational analysis demonstrates that the optimal policy increases profit significantly, particularly when demand has high seasonality and when production capacity is tight.

2.6 Appendix A

Proof of Lemma 2.1

Proof: By contradiction, assume that there is an optimal policy with $R_t \cdot B_t > 0$ for some period t. Let $\bar{R}_t = R_t - 1$ and $\bar{B}_t = B_t - 1$ be the alternative policy. Let C_t and $\bar{C}_t$ be the cost incurred in period t under the two policies respectively, and let I_t and $\bar{I}_t$ be the inventory level at the beginning of period $t+1$. We compare the two policies in the following four cases:

- Case 1: $Y_t \leq 0$, then $R_t = 0$, and the lemma holds.
- Case 2: $Y_t > 0$ and $D_t \leq Y_t - R_t$, hence $D_t < Y_t - \bar{R}_t$.

  We have $C_t = h_t(Y_t - D_t)$, $I_{t+1} = Y_t - D_t$, $\bar{C}_t = h_t(Y_t - D_t)$, $\bar{I}_{t+1} = Y_t - D_t$. So $\bar{C}_t = C_t, \bar{I}_{t+1} = I_{t+1}$.
- Case 3: $Y_t > 0$ and $D_t > Y_t - R_t$, hence $D_t \geq Y_t - \bar{R}_t$.

$$\begin{aligned}
C_t &= h_t R_t + b_t \min(B_t, D_t - Y_t + R_t) + r_t(D_t - Y_t + R_t - B_t)^+ \\
I_{t+1} &= R_t - \min(B_t, D_t - Y_t + R_t) \\
\bar{C}_t &= h_t \bar{R}_t + b_t \min(\bar{B}_t, D_t - Y_t + \bar{R}_t) + r_t(D_t - Y_t + \bar{R}_t - \bar{B}_t)^+ = C_t - h_t - l_t < C_t \\
\bar{I}_{t+1} &= \bar{R}_t - \min(\bar{B}_t, D_t - Y_t + \bar{R}_t) = I_{t+1}
\end{aligned}$$

With the alternative policy, the selling revenue is the same, and the inventory left to the future is also the same, but the cost is less than or equal to the original policy. So the alternative policy is always better than the original policy, which contradicts that the original policy is optimal.

The following lemma is used in the proof for Theorem 2.1.

Lemma 2.2 *Given $g(x, y)$ is jointly concave in x and y, $G(x) = \max_y g(x, y)$ is a concave function for x.*

Proof: For any $x_1, x_2 \in R$, let $y_1 = \arg\max\{y | g(x_1, y)\}$, $y_2 = \arg\max\{y | g(x_2, y)\}$. For any $\lambda \in [0, 1]$, let $x_\lambda = \lambda x_1 + (1 - \lambda)x_2$, $y_\lambda = \lambda y_1 + (1 - \lambda)y_2$. We have $G(x_\lambda) = \max_y g(x_\lambda, y) \geq g(x_\lambda, y_\lambda) \geq \lambda g(x_1, y_1) + (1 - \lambda)g(x_2, y_2) = \lambda G(x_1) + (1 - \lambda)G(x_2)$.

Proof of Concavity in Delayed Production Strategy (Theorem 2.1)

Proof: We prove by induction.

1. For period $t = T$, we have $B_T = 0$, $R_T = 0$ and $J_{T+1}(I_T) = vI_T$.

$$G_T^R(Y_T) = \int_Y^\infty [p_T Y_T - r_T(D_T - Y_T)]d\Phi_T + \int_0^Y [p_T D_T - h_T(Y_T - D_T) + v(Y_T - D_T)]d\Phi_T$$
$$G_T^B(Y_T) = \int_Y^\infty [p_T Y_T - r_T(D_T - Y_T)]d\Phi_T + \int_0^Y [p_T D_T - h_T(Y_T - D_T) + v(Y_T - D_T)]d\Phi_T$$

So $G_T(Y_T) = G_T^R(Y_T) = G_T^B(Y_T)$, and $G_T'(Y_T) = \int_{Y_T}^\infty (p_T + r_T)d\Phi_T + \int_0^{Y_T}(v - h_t)d\Phi_T$. Thus $G_T(Y_T)$ is concave in Y_T since $G_T'(Y_T)$ is non-increasing in Y_T: $G_T''(Y_T) = (v - h_T - p_T - r_T)\phi_T(Y_T) \leq 0$.

2. Let $j_t(I_t, Y_t) = -c_t(Y_t - I_t) + G_t(Y_t)$, so $J_t(I_t) = \max_{Y_t: I_t \leq Y_t \leq I_t + q_t} j_t(I_t, Y_t)$. Given t, $t = 1, ..., T$, assume that $G_{t+1}(Y_{t+1})$ is concave in Y_t, then we can prove that $j_{t+1}(I_{t+1}, Y_{t+1})$ is jointly concave in I_{t+1} and Y_{t+1} by the following. For any $(I_1, Y_1), (I_2, Y_2)$, let $I_\lambda = \lambda I_1 + (1-\lambda)I_2$, $Y_\lambda = \lambda Y_1 + (1-\lambda)Y_2$. Then,

$$\begin{aligned}
j_{t+1}(I_\lambda, Y_\lambda) &= -c_{t+1}(Y_\lambda - I_\lambda) + G_{t+1}(Y_\lambda) \\
&= -c_{t+1}(\lambda Y_1 + (1-\lambda)Y_2 - \lambda I_1 - (1-\lambda)I_2) + G_{t+1}(\lambda Y_1 + (1-\lambda)Y_2) \\
&\geq -\lambda c_{t+1}(Y_1 - I_1) - (1-\lambda)c_{t+1}(Y_2 - I_2) + \lambda G_{t+1}(Y_1) + (1-\lambda)G_{t+1}(Y_2) \\
&= \lambda j_{t+1}(I_1, Y_1) + (1-\lambda)j_{t+1}(I_2, Y_2).
\end{aligned}$$

So by Lemma 2, $J_{t+1}(I_{t+1})$ is concave in I_{t+1}, and as a result $J_{t+1}'(I_{t+1})$ is non-increasing in I_{t+1}.

3. Next, we are going to prove the quasi-concavity of $g_t^R(Y_t, R_t)$ with respect to R_t. First let us examine the derivative of $g_t^R(Y_t, R_t)$ with respect to R_t.

$$g_t'^R(R_t) = \begin{cases} 0 & if\ Y_t < R_t \\ (-p_t - r_t - h_t + J_{t+1}'(R_t))(1 - \Phi_t(Y_t - R_t)) & if\ otherwise \end{cases}$$

Let us define R_t^* as

$$R_t^* = \begin{cases} \max\{I : p_t + r_t + h_t < J_{t+1}'(I)\} & if\ p_t + r_t + h_t < J_{t+1}'(0) \\ 0 & otherwise. \end{cases}$$

Thus we have $g_t'^R(R_t) \geq 0$, when $0 \leq R_t \leq R_t^*$, and $g_t'^R(R_t) \leq 0$, when $R_t > R_t^*$, so, $g_t^R(Y_t, R_t)$ is quasi-concave with respect to R_t. R_t^* is the unique unconstrained optimizer of $g_t^R(Y_t, R_t)$ and it is independent of inventory level Y_t. $R_t^c = \min(R_t^*, Y_t)$ maximizes $g_t^R(Y_t, R_t)$, for $0 \le R_t \le Y_t$.

4. Next let us prove the quasi-convexity of $g_t^B(Y_t, B_t)$ with respect to B_t. First let us examine the derivative of $g_t^B(Y_t, B_t)$ with respect to B_t.

$$g_t^{'B}(B_t) = \begin{cases} 0 & if\ Y_t^- > B_t \\ (p_t + r_t - b_t - J_{t+1}^{'}(-B_t))(1 - \Phi_t(Y_t + R_t)) & if\ Y_t^- \le B_t \end{cases}$$

Let us define B_t^* as

$$B_t^* = \begin{cases} \max\{I : p_t + r_t - b_t \ge J_{t+1}^{'}(-I)\} & if\ p_t + r_t - b_t \ge J_{t+1}^{'}(0) \\ 0 & otherwise. \end{cases}$$

Thus we have $g_t^{'B}(B_t) \ge 0$, when $0 \le B_t \le B_t^*$, and $g_t^{'B}(B_t) \le 0$, when $B_t > B_t^*$, so, $g_t^B(Y_t, B_t)$ is quasi-concave with respect to B_t. B_t^* is the unique unconstrained optimizer of $g_t^B(Y_t, B_t)$, and it is independent of inventory level Y_t. $B_t^c = \max(\min(B_t^*, \sum_{i=1}^N q_{t+i}), Y_t^-)$ maximizes $g_t^B(Y_t, B_t)$, for $0 \le B_t \le q_{t+1}$.

5. Let us prove the concavity of $G_t^R(Y_t)$ with respect to Y_t.

   – First let us examine the derivative of $G_t^R(Y_t)$ with respect to Y_t.

$$\begin{aligned} G_t^R(Y_t) = g_t^R(Y_t, R_t^c) = & \int\{p_t \min(D_t, Y_t - R_t^c) - h_t \max(R_t^c, Y_t - D_t) \\ & -r_t(D_t - Y_t + R_t^c)^+ + J_{t+1}(\max(R_t^c, Y_t - D_t))\}d\Phi_t. \end{aligned}$$

$$G_t^{'R}(Y_t) = \begin{cases} J_{t+1}'(Y_t) & if\ Y_t < 0 \\ J_{t+1}'(Y_t) - h_t & if\ 0 \le Y_t \le R_t^* \\ \int_{Y_t - R_t^*}^{\infty}(p_t + r_t)d\Phi_t + \int_0^{Y_t - R_t^*}(J_{t+1}'(Y_t - D_t) - h_t)d\Phi_t & if\ Y_t > R_t^*. \end{cases}$$

   – We consider $G_t^{''R}(Y_t)$ in five cases:

   (a)Case 1: $Y_t < 0$:

   $G_t^{''R}(Y_t) = J_{t+1}''(Y_t) \le 0$, due to the concavity of J_{t+1}.

   (b)Case 2: $Y_t = 0$:

   $G_t^{''R}(Y_t) = J_{t+1}''(Y_t) - h_t \le 0$, due to the concavity of J_{t+1}.

   (c)Case 3: $0 < Y_t < R_t^*$:

   $G_t^{''R}(Y_t) = J_{t+1}''(Y_t) \le 0$, due to the concavity of J_{t+1}.

   (d)Case 4: $Y_t = R_t^*$:

$G_t''^R(Y_t) = p_t + r_t + h_t - J'_{t+1}(R_t^*) \le 0$

due to the choice of R_t^* and the concavity of J_{t+1}.

(e)Case 5: $Y_t > R_t^*$:

$G_t''^R(Y_t) = \int_0^{Y_t - R_t^*} J''_{t+1}(Y_t - D_t)d\Phi_t + (J'_{t+1}(R_t^*) - p_t - r_t - h_t)\phi_t(Y_t - R_t^*) \le 0$

- Since $G_t''^R(Y_t) \le 0$ for all Y_t, $G_t^R(Y_t)$ is concave in Y_t.

6. Let us prove the concavity of $G_t^B(Y_t)$ with respect to Y_t.

   - First let us examine the derivative of $G_t^B(Y_t)$ with respect to Y_t.

$$\begin{aligned} G_t^B(Y_t) = g_t^B(Y_t, B_t^*) = \; & \int\{p_t \min(D_t, Y_t + B_t^*) - h_t(Y_t - D_t)^+ - r_t(D_t - Y_t - B_t^c)^+ \\ & -b_t \min((D_t - Y_t)^+, B_t^c) + J_{t+1}(\max(Y_t - D_t, -B_t^c))\}\Phi_t \end{aligned}$$

$$G_t'^B(Y_t) = \begin{cases} \int_0^{Y_t}[-h_t + J'_{t+1}(Y_t - D_t)]d\Phi_t + \int_{Y_t}^{Y_t+B_t^*}[l_t + J'_{t+1}(Y_t - D_t)]d\Phi_t & \\ + \int_{Y_t+B_t^*}^{\infty}[p_t + r_t]d\Phi_t & if\ Y_t \ge 0 \\ \int_0^{Y_t+B_t^*}[l_t + J'_{t+1}(Y_t - D_t)]d\Phi_t + \int_{Y_t+B_t^*}^{\infty}[p_t + r_t]d\Phi_t & if\ -B_t^* \le Y_t < 0 \\ l_t + J'_{t+1}(Y_t - D_t) & if\ Y_t < B_t^*. \end{cases}$$

   - We consider $G_t''^B(Y_t)$ in four cases:

     (a)Case 1: $Y_t \ge 0$:

     $G_t''^B(Y_t) = \int_0^{Y_t+B_t^*} J''_{t+1}(Y_t - D_t)d\Phi_t + \phi_t(Y_t + B_t^*)[-p_t - r_t + l_t + J'_{t+1}(-B_t^*)] +$

     $\phi_t(Y_t)(-h_t - l_t) \le 0$, due to the concavity of J_{t+1} and the choice of B_t^*.

     (b)Case 2: $-B_t^* < Y_t < 0$:

     $G_t''^B(Y_t) = \int_0^{Y_t+B_t^*} J''_{t+1}(Y_t - D_t)d\Phi_t + \phi_t(Y_t + B_t^*)[-p_t - r_t + l_t + J'_{t+1}(-B_t^*)] \le 0,$

     due to the concavity of J_{t+1} and the choice of B_t^*.

     (c)Case 3: $Y_t = -B_t^*$:

     $G_t''^B(Y_t) = -p_t - r_t + l_t + J'_{t+1}(-B_t^*)$, due to the concavity of J_{t+1} and the choice of B_t^*.

     (d)Case 4: $Y_t < -B_t^*$:

     $G_t''^R(Y_t) = J''_{t+1}(Y_t) \le 0$

     due to the choice of R_t^* and the concavity of J_{t+1}.

   - Since $G_t''^B(Y_t) \le 0$ for all Y_t, $G_t^B(Y_t)$ is concave in Y_t.

7. Let us prove the concavity of $G_t(Y_t)$.

   In each period, we must be in one of the following cases, which are independent of the value of Y_t:

   - If $(p_t + r_t - b_t) \leq J'_{t+1}(0) \leq (p_t + r_t + h_t)$, we have $R_t^* = B_t^* = 0$, therefore $R_t^c = B_t^c = 0$, so we have $G_t(Y_t) = G_t^R(Y_t) = G_t^B(Y_t)$, which can be seen from the formulations directly.
   - If $J'_{t+1}(0) > (p_t + r_t + h_t)$ and $Y_t \geq 0$ we have $R_t^* \geq 0$ and $B_t^* = 0$, therefore $R_t^c \geq 0$ and $B_t^c = 0$, so we have $G_t(Y_t) = G_t^R(Y_t) \geq G_t^B(Y_t)$.
   - If $J'_{t+1}(0) > (p_t + r_t + h_t)$ and $Y_t < 0$ we have $R_t^* = 0$ and $B_t^* = 0$, therefore $R_t^c \geq 0$ and $B_t^c = 0$, so we have $G_t(Y_t) = G_t^R(Y_t) = G_t^B(Y_t)$.
   - If $(p_t + r_t - b_t) < J'_{t+1}(0)$, we have $B_t^* \geq 0$ and $R_t^* = 0$, therefore $B_t^c \geq 0$ and $R_t^c = 0$, so we have $G_t(Y_t) = G_t^B(Y_t) \geq G_t^R(Y_t)$.

   We see that in each period, $G_t(Y_t)$ reduces to some function that is proved to be concave. Therefore $G_t(Y_t)$ is concave.

2.7 Appendix B

Experimental Data

Periods	**c**	p	$E(D)$
1	80	100	100
2	90	110	80
3	80	100	100
4	70	90	120
5	80	100	100
6	90	110	80
7	80	100	100
8	70	90	120
9	80	100	100
10	90	110	80
11	80	100	100
12	70	90	120
Avg	**80**	**100**	**100**

Table 2.1: Experimental Data for Single-Class-Customer Problem

Chapter 3

Two-Class-Customer Problem with Backlogged and Discretionary Sales

3.1 Introduction

In this chapter, we propose an approach to practice market segmentation in the manufacturing industry by differentiating customers according to their sensitivity to lead time. For example, many manufacturing companies face the following problem: some customers are willing to pay high price, but they are not willing to wait for delayed fulfillments; other customers are willing to wait for delayed fulfillments, but they can only pay low prices. Since the manufacturer has a limited production capacity, the manufacturer needs to determine how to allocate the capacity effectively in order to maximize the profit. If the manufacturer can distinguish between different classes, then the manufacturer can serve different classes with different prices and different lead times, and manage production and inventory appropriately. On the other hand, if the manufacturer cannot differentiate the customers, s/he must serve them as a single class. We call the former strategy the Differentiation Strategy, the latter the Nondifferentiation Strategy, and compare them with the same exogenous two-class demands to examine the effect of market segmentation.

The rest of this chapter is organized as follows. In Section 3.2, we present the notations and assumptions for both strategies. We study the Nondifferentiation Strategy in Section 3.3, and the Differentiation Strategy in Section 3.4. In Section 3.5, we do computational analysis to compare the manufacturer's expected profits under these two strategies, and we examine the effectiveness of the market segmentation in the manufacturing industry.

3.2 Assumptions and Notations

We analyze a model with a single manufacture, a single type of product and two classes of customers. We assume that the 1^{st} class customers would not wait for delayed fulfillments (lost sales), and the 2^{nd} class customers are willing to wait for one period for delayed fulfillments (partial backorders). We assume that the amount of demands in each class is a non-stationary, time-dependent, general stochastic function, D_t^i, but we do not assume a particular distribution. We continue to use the notations in Chapter 2, and we add a class index superscript to the parameters and variables associated with each demand class.

We study both the Differentiation Strategy and the Nondifferentiation Strategy. Under the Nondifferentiation Strategy, the sequence of events in every period is as follows: at the beginning of a period, the manufacturer checks the inventory level and decides on the production quantity, the minimum amount of inventory to reserve for future sales, and the maximum amount of backorders to be fulfilled in the next period; products arrive in zero lead time, and the manufacturer fulfills the backorders with the available inventory; and as demands arrive during the period, the manufacturer deals all the demands as from a single class, and the manufacturer realizes, backlogs, or rejects demands with respect to the FCFS (first come fist serve) policy.

Under the Differentiation Strategy, the sequence of events at the beginning of each period is exactly the same as under the Non-differentiation Strategy. During the period, the manufacturer will use different policies to different classes of customers. The manufacturer will satisfy up to $Y_t - R_t$ units of demands from the 1^{st} class and backlog up to B_t units of demands from the 2^{nd} class. The 2^{nd} class demands will not be realized immediately even if the manufacturer has enough inventories.

In the following, we study the models under two strategies respectively.

3.3 Nondifferentiation Strategy

3.3.1 Model

In this section, we study the Nondifferentiation Strategy, which is an extension of the optimal policy for the single-class-customer problem in Chapter 2. This model is motivated by the reality that some manufacturers are not able to treat customers differently due to some legislative or industry constraints, even though they know there are multiple classes of customers.

We assume that under the Nondifferentiation Strategy, the manufacturer serves the customers as a single class. The manufacturer takes the 2^{nd} class customers reservation price, p_t^2, as the selling price to all customers. So the total amount of demand in period t is $D_t = D_t^1 + D_t^2$. Among all the customers, only those from the 2^{nd} class are willing to wait for delayed fulfillments, and we let α_t be the proportion of demand ordered by the 2^{nd} class customers, $\alpha_t = E[D_t^2]/E[D_t^1 + D_t^2]$. Let r be the weighted average rejection penalty for unsatisfied demands, $r_t = (1-\alpha_t)r_t^1 + \alpha_t r_t^2$. Given a price vector, the profit-to-go function under the Nondifferentiation Strategy is

$$J_t(I_t) = \max_{Y_t:\max(0,I_t)\leq Y_t\leq I_t+q_t} \{-c_t(Y_t - I_t) + G_t(Y_t)\}. \tag{3.1}$$

The first term in 3.1is the production cost, and the second term is the profit-to-go with Y_t units of products available after production but before satisfying new demands. $G_t(Y_t)$ can be calculated as,

$$\begin{aligned} G_t(Y_t) = \max_{0\leq R_t\leq Y_t, 0\leq B_t\leq q_{t+1}} \quad & \int\{p_t^2 \min(D_t, Y_t - R_t + \min(B_t, \alpha(D_t - Y_t + R_t)^+)) \\ & -h_t \max(R_t, Y_t - D_t) - r_t^1(1-\alpha)\min(B_t/\alpha, (D_t - Y_t + R_t)^+)) \\ & -r_t(D_t - Y_t + R_t - B_t/\alpha)^+ - b_t^2 \min(B_t, \alpha(D_t - Y_t + R_t)^+) \\ & +J_{t+1}(\max(R_t, Y_t - D_t) - \min(\alpha(D_t - Y_t)^+, B_t))\}d\Phi_t. \end{aligned} \tag{3.2}$$

Finally, we let $J_{T+1}(I_T) = v \cdot I_T$, which is the expected salvage value of leftover inventory.

The first term in (3.2) is the selling revenue. The second term is the inventory holding cost. The third term is the penalty associated with the lost 1^{st} class demands who are not willing to be backlogged. The fourth term is the rejection penalty for demands beyond the acceptance level. The fifth term is the delay penalty associated with the backlogged 2^{nd} class demands. The last term represents the profit-to-go from the end of this period.

Next, we show that in an optimal policy for the Non-Differentiation Strategy, in any period, either the amount of reserved inventory equals zero or the amount of backlogged demands equals zero.

Lemma 3.1 *In any optimal policy under the Nondifferentiation Strategy, we have* $R_t \cdot B_t = 0$, *for* $t = 1, 2, ..., T$.

Please refer to the appendix for the proof. The intuition behind the lemma is very simple. Suppose that $R_t > 0$, which means that the manufacturer may reject some high-price demands in period t in order to reserve some inventory to period $t+1$, then it would not be profitable for the manufacturer to use the inventory in period $t+1$ to fulfill any low-price demand in period t, thus we will have $B_t = 0$. The intuitive explanation for the case with $B_t > 0$ is similar.

With Lemma 3.1, the structure of the optimal policies can be simplified. Under the Nondifferentiation Strategy, the manufacturer choose one out of the two policies: either the reserve-inventory policy ($R_t \geq 0, B_t = 0$) or the backlog-demand policy ($B_t \geq 0, R_t = 0$), thus,

$$G_t(Y_t) = \max\{G_t^R(Y_t), G_t^B(Y_t)\}.$$

$G_t^R(Y_t)$ and $G_t^B(Y_t)$ are calculated as:

$$\begin{aligned} G_t^R(Y_t) = & \max_{R_t: 0 \leq R_t \leq Y_t} g_t^R(Y_t, R_t), \\ G_t^B(Y_t) = & \max_{B_t: 0 \leq B_t \leq q_{t+1}} g_t^B(Y_t, B_t), \end{aligned} \tag{3.3}$$

where $g_t^R(Y_t, R_t)$ and $g_t^B(Y_t, B_t)$ are:

$$
\begin{aligned}
g_t^R(Y_t, R_t) = \;& \int \{p_t^2 \min(D_t, Y_t - R_t) - h_t \max(R_t, Y_t - D_t) - r_t(D_t - Y_t + R_t)^+ \\
& + J_{t+1}(\max(R_t, Y_t - D_t))\} d\Phi_t,
\end{aligned}
\tag{3.4}
$$

$$
\begin{aligned}
g_t^B(Y_t, B_t) = \;& \int \{p_t^2 \min(D_t, Y_t + \min(B_t, \alpha(D_t - Y_t)^+)) - h_t(Y_t - D_t)^+ \\
& - r_t^1(1-\alpha)\min(B_t/\alpha, (D_t - Y_t)^+)) - r_t(D_t - Y_t - B_t/\alpha)^+ \\
& - b_t^2 \min(B_t, \alpha(D_t - Y_t)^+) + J_{t+1}((Y_t - D_t)^+ - \min(\alpha(D_t - Y_t)^+, B_t))\} d\Phi_t.
\end{aligned}
\tag{3.5}
$$

3.3.2 Optimal Policy

Under the Nondifferentiation Strategy, we show that functions $g_t^R(Y_t, R_t)$ and $g_t^B(Y_t, B_t)$ are quasi-concave in R_t and B_t respectively, and each of them has a unique unconstrained optimizer that is independent of the inventory level Y_t. We also show that the expected profit-to-go functions, $J_t(I_t)$ and $G_t(Y_t)$, are concave in I_t and Y_t respectively. These results are summarized in the following theorem.

Theorem 3.1 *Under the Nondifferentiation Strategy,*

- $g_t^R(Y_t, R_t)$ *is a quasi-concave function of* R_t*, for all* $t = 1, ..., T$.
- $g_t^B(Y_t, B_t)$ *is a quasi-concave function of* B_t*, for all* $t = 1, ..., T$.
- $G_t(Y_t)$ *is a concave function of* Y_t*, for all* $t = 1, ..., T$.
- $J_t(I_t)$ *is a concave function of* I_t*, for all* $t = 1, ..., T$.
- *The unconstrained optimizers for* $g_t^R(Y_t, R_t)$ *and* $g_t^B(Y_t, B_t)$*,* R_t^* *and* B_t^**, are independent of inventory level* Y_t*, where*

$$R_t^*(Y_t) = \arg\max_{0 \le R_t}\{g_t^R(Y_t, R_t)\}, B_t^*(Y_t) = \arg\max_{0 \le B_t}\{g_t^B(Y_t, B_t)\}. \tag{3.6}$$

Let R_t^c and B_t^c be the constrained optimizers of $g_t^R(Y_t, R_t)$ and $g_t^B(Y_t, B_t)$, $R_t^c = \min(R_t^*, (Y_t)^+)$, $B_t^c = \min(B_t^*, q_{t+1})$. Theorem 3.1 implies the optimal policy for the Nondifferentiation Strategy, and thus we have the following corollary.

Corollary 3.1 *Given a vector of prices, there exists an optimal policy for the Nondifferentiation Strategy with an optimal order-up-to level (S_t^*), an optimal reserve-up-to-level (R_t^c) and an optimal backlog-up-to level (B_t^c).*

Thus the Nondifferentiation Strategy leads to an optimal policy characterized by three parameters, and we denote it as the (S, R, B) policy. For the order-up-to level, the policy is to produce to bring the inventory level up to S_t^c. If the reserve-up-to level is positive, the manufacturer will set aside R_t^c units of inventory for the future, he will only accept $Y_t - R_t^c$ units of demands; if the backlog-up-to level is positive, the manufacturer will satisfy up to Y_t units of demands and backlog up to B_t^c units of demands and satisfy them in the next period.

3.4 Differentiation Strategy

3.4.1 Model

Under the Differentiation Strategy, the manufacturer offers different prices and different lead times to two classes of customers. Given two price vectors, the profit-to-go function under the Differentiation Strategy is:

$$J_t(I_t) = \max_{Y_t:\max\{0,I_t\}\leq Y_t\leq I_t+q_t} \{-c_t(Y_t - I_t) + G_t(Y_t)\}, \tag{3.7}$$

where the first term is the production cost, and the second term is the profit-to-go with Y_t units of products available after production but before satisfying new demands. $G_t(Y_t)$ is defined as:

$$\begin{aligned} G_t(Y_t) = \max_{0\leq R_t\leq Y_t, 0\leq B_t\leq q_{t+1}} \quad & \int\int\{p_t^1 \min(D_t^1, Y_t - R_t) - h_t \max(R_t, Y_t - D_t^1) \\ & -r_t^1(D_t^1 - Y_t + R_t)^+ + (p_t^2 - b_t^2)\min(D_t^2, B_t) \\ & -r_t^2(D_t^2 - B_t)^+ + J_{t+1}((Y_t - D_t^1)^+ - \min(D_t^2, B_t))\}d\Phi_t^1 d\Phi_t^2 \end{aligned} \tag{3.8}$$

Finally, we let $J_{T+1}(I_T) = v \cdot I_T$, which is the expected salvage value of leftover inventory.

The first term in (3.8) is the selling revenue to the 1^{st} class. The second term is the inventory holding cost. The third term is the penalty associated with rejected demands of the 1^{st} class. The fourth term is the selling revenue to the 2^{nd} class. The fifth term is the penalty associated with rejected demands of the 2^{nd} class. The sixth term represents the profit-to-go from the end of this period.

Next, we show that in an optimal policy for the Differentiation Strategy, in any period, either the amount of reserved inventory equals zero or the amount of backlogged demands equals zero.

Lemma 3.2 *In any optimal policy under the Differentiation Strategy, we have $R_t \cdot B_t = 0$, for $t = 1, 2, ..., T$.*

Please refer to the appendix for the proof. With Lemma 3.2, the structure of the optimal policies can be simplified. The manufacturer can choose one out of the two policies: either the reserve-inventory policy ($R_t \geq 0, B_t = 0$) or the backlog-demand policy ($B_t \geq 0, R_t = 0$), thus,

$$G_t(Y_t) = \max\{G_t^R(Y_t), G_t^B(Y_t)\}.$$

$G_t^R(Y_t)$ and $G_t^B(Y_t)$ represent the profit-to-go with Y_t units of products available after production under the reserve-inventory policy or the backlog-demand policy respectively, and they are given by,

$$G_t^R(Y_t) = \max_{R_t : 0 \leq R_t \leq Y_t} g_t^R(Y_t, R_t) \tag{3.9}$$

$$G_t^B(Y_t) = \max_{B_t : 0 \leq B_t \leq q_{t+1}} g_t^B(Y_t, B_t) \tag{3.10}$$

where $g_t^R(Y_t, R_t)$ indicates the profit-to-go with Y_t units of products available after production and R_t units are reserved for the next period under the reserve-inventory policy; $g_t^B(Y_t, B_t)$ indicates the profit-to-go with Y_t units of products available after production and at most another B_t units of demands can be backlogged under the backlog-demand policy. Thus,

$$\begin{aligned} g_t^R(Y_t, R_t) = \; & \int\int\{p_t^1 \min(D_t^1, Y_t - R_t) - h_t \max(R_t, Y_t - D_t^1) - r_t^1(D_t^1 - Y_t + R_t)^+ \\ & -r_t^2 D_t^2 + J_{t+1}(\max(R_t, Y_t - D_t^1))\} d\Phi_t^1 d\Phi_t^2 \end{aligned} \tag{3.11}$$

$$g_t^B(Y_t, B_t) = \int\int\{p_t^1 \min(D_t^1, Y_t) - h_t(Y_t - D_t^1)^+ - r_t^1(D_t^1 - Y_t)^+ - r_t^2(D_t^2 - B_t)^+ \\ +(p_t^2 - b_t^2)\min(D_t^2, B_t) + J_{t+1}((Y_t - D_t^1)^+ - \min(D_t^2, B_t))\}d\Phi_t^1 d\Phi_t^2 \tag{3.12}$$

3.4.2 Structural Results

In the following two sections, we analyze the properties of the profit-to-go functions in period t, $g_t^R(Y_t, R_t)$ and $g_t^B(Y_t, B_t)$, under the condition that the profit-to-go function in period $t+1$, $J_{t+1}(I_{t+1})$, is a concave function.

Condition 3.1 *$J_{t+1}(I_{t+1})$ is concave in I_{t+1}.*

Under this condition, for a given value of Y_t, both the profit-to-go functions of $g_t^R(Y_t, R_t)$ and $g_t^B(Y_t, B_t)$ are quasi-concave in R_t and B_t respectively.

Lemma 3.3 *For a given value of Y_t, $g_t^R(Y_t, R_t)$ is a quasi-concave function of R_t.*

Lemma 3.4 *For a given value of Y_t, $g_t^B(Y_t, B_t)$ is a quasi-concave function of B_t.*

Due to the quasi-concavity, for a given value of Y_t, there exist unique maximizers for $g_t^R(Y_t, R_t)$ and $g_t^B(Y_t, B_t)$ respectively.

$$R_t^*(Y_t) = \max\{R | \frac{\partial g_t^R(Y_t, R_t)}{\partial R_t} \geq 0\}$$
$$B_t^*(Y_t) = \max\{B | \frac{\partial g_t^B(Y_t, B_t)}{\partial B_t} \geq 0\}$$

Then we will show that the un-constrained optimizer for $g_t^R(Y_t, R_t)$, $R_t^*(Y_t)$, is independent of Y_t, and $G_t^R(Y_t)$ is concave in Y_t.

Property 3.1 *$R_t^*(Y_t)$ is independent of Y_t.*

Property 3.2 *$G_t^R(Y_t)$ is a concave function of Y_t.*

These properties greatly simplify the structure of the reserve-inventory policy.

Corollary 3.2 *The optimal reserve-inventory policy is characterized with an optimal order-up-to level (S_t^*) and an optimal reserve-up-to level (R_t^*).*

The optimal reserve-inventory policy is to produce enough to bring the inventory level up to S_t^* if there is sufficient capacity, otherwise produce to the maximum capacity. The manufacturer will set aside up to R_t^* units of inventory for the future, he will only accept $(Y_t - R_t^*)^+$ units of demands from the 1^{st} class, and he will reject all the demands from the 2^{nd} class. We denote (S_t^c, R_t^c) as the constrained optimizers for $g_t^R(Y_t, R_t)$.

$$S_t^c = \min(S_t^*, I_t + q_t) \qquad R_t^c = \min(R_t^*, Y_t)$$

3.4.3 Heuristic Analysis

Under the backlog-demand policy, the un-constrained optimizer for $g_t^B(Y_t, B_t)$, $B_t^*(Y_t)$, depends on Y_t, which makes the the control policy under the backlog-demand policy more complicated than that under the reserve-inventory policy. Let us denote,

$$\underline{B}_t = \begin{cases} \min\{I : J_{t+1}'(-I) \geq p_t^2 + \beta_t^2 - \ell_t^2\} & if \quad J_t'(0) \leq p_t^2 + \beta_t^2 - \ell_t^2 \\ 0 & if \quad O.W. \end{cases} \tag{3.13}$$

$$\overline{B}_t = \begin{cases} \min\{I : J_{t+1}'(-I) \geq p_t^1 + \beta_t^1 + h_t^1\} & if \quad J_t'(0) \leq p_t^1 + \beta_t^1 + h_t^1 \\ 0 & if \quad O.W. \end{cases} \tag{3.14}$$

It is easy to show that $B_t^*(Y_t) \geq \underline{B}_t$.

Property 3.3 $B_t^*(Y_t) \geq \underline{B}_t$, *for* $Y_t \geq 0$.

On the other hand, $B_t^*(Y_t)$ could be either greater or smaller than $\overline{B}_t$. In the following, we discuss properties of function $g_t^B(B_t, Y_t)$ under the condition of $B_t^*(Y_t) \leq \overline{B}_t$ to derive a heuristic policy.

Condition 3.2 $B_t^*(Y_t) \leq \overline{B}_t$, *for* $Y_t \geq 0$.

The following lemma indicates that if Condition 3.2 is satisfied, then there exists a unique pair of (B_t^*, Y_t^*) maximizing $g_t^B(B_t, Y_t) - c_t(Y_t - I_t)$.

Lemma 3.5 *If Condition 3.2 is satisfied,* $g_t^B(B_t, Y_t) - c_t(Y_t - I_t)$ *is jointly quasi-concave in* B_t *and* Y_t *for* $B_t \in [0, B_t^*(Y_t)]$.

We can maximize $g_t^B(B_t, Y_t) - c_t(Y_t - I_t)$ over Y_t and B_t together. We denote S_t^c and B_t^c as the constrained optimizers of $g_t^B(Y_t, B_t)$.

$$(S_t^c, B_t^c) = \underset{(Y_t, B_t)}{\arg} \max\{g_t^B(Y_t, B_t) - c_t Y_t | 0 \leq Y_t \leq I_t + q_t, 0 \leq B_t \leq \min\{q_{t+1}, \overline{B}_t\}\}$$

Lemma 3.5 implies the optimal policy under Condition 3.2 for the backlog-demand policy, and thus we have the following corollary.

Corollary 3.3 *If Condition 3.2 is satisfied, the optimal backlog-demand policy is characterized with an optimal order-up-to level (S_t^c) and an optimal backlog-up-to level (B_t^c).*

The optimal reserve-inventory policy is to produce to bring the inventory level up to S_t^c. The manufacturer will satisfy up to S_t^c units of demands from the 1^{st} class, and he will accept up to B_t^c units of demands from the 2^{nd} class.

Under Condition 3.2, we could further explore properties for the backlog-demand policy.

Property 3.4 *If Condition 3.2 is satisfied, $G_t^B(Y_t)$ is concave in Y_t.*

Property 3.5 *If Condition 3.2 is satisfied, $G_t(Y_t)$ is concave in Y_t.*

Property 3.6 *If Condition 3.2 is satisfied, $J_t(I_t)$ is concave in I_t.*

Later, in the numerical analysis, we will show that Condition 3.2 holds in 93% of the cases we examined, and it is 100% satisfied if the proportion of first class demand is greater than 30%. If Condition 3.2 is not satisfied, i.e., $B_t^*(Y_t) > \overline{B}_t$, the joint quasi-concavity may not hold, and we use $\overline{B}_t$ instead of $B_t^*(Y_t)$, and we use it as the heuristic policy.

In summary, the Differentiation Strategy leads to a heuristic policy characterized by three parameters, S^c, R^c, and B^c. For the order-up-to level, the policy is to produce to bring the inventory level up to S_t^c. If the reserve-up-to level is positive, the manufacturer will set aside R_t^c units of inventory for the future, he will only accept $Y_t - R_t^c$ units of demands from the 1^{st} class, and he will reject all the demands from the 2^{nd} class; if the backlog-up-to level is positive, the manufacturer will satisfy up to Y_t units of demands from the 1^{st} class and backlog up to B_t^c units of demands from the 2^{nd} class.

3.5 Computational Analysis

In this chapter, we have analyzed two strategies: the Differentiation Strategy (DS) and the Nondifferentiation Strategy (NS). In the following, we report a computational study conducted to obtain insights about the benefits of these strategies. Our goal is to examine the relative performance of the (S, R, B) policies and identify the situations where the (S, R, B) policies can provide significant profit increase.

The Differentiation Strategy improves the profit in two ways: market segmentation and shifting inventory to better match demand. The Nondifferentiation Strategy, on the other hand, only shifts inventory. So comparing the Differentiation Strategy with the Nondifferentiation Strategy, we can quantify the profit improvement contributed by each factor.

The benchmark we use for each of our strategies is a traditional base-stock policy, namely, when the manufacturer uses the modified order-up-to policy (S policy) and serves all the customers as in a single class. In the benchmark policy, unsatisfied demands are lost if the inventory is not available. We compare this traditional policy to each of our (S, R, B) policies using the benchmark ratio. We define the benchmark ratio, *Profit Potential*, as

$$Profit\ Potential = \frac{V_{(S,R,B)}}{V_S} - 1, \tag{3.15}$$

where V indicates the expected profit of the problem being solved. In the traditional policy, we use p_t^2 as the price charged to customers.

For demand variability, we focus on demand uncertainty, which is assumed to be additive to demand with a mean of 0. We define the coefficient of variation of demand uncertainty in a given period as $CV_U^i = s(D_t^i)/E(D_t^i)$, where s denotes the standard deviation, E denotes the expected value and the index $i = 1, 2$ represents the demand class. In all cases shown, the coefficient variation of demand uncertainty is the same in each period and equals 0.2.

Production capacity is constant for a particular instance, while it is allowed to take the values of 60% (low), 80% (med), and 100% (high) of the expected average demand over the horizon, indicated as Dem^*. In the numerical study, the total average demand equals 100 in each experiment. The production cost is the same across instances but varies by period. Similar experimental results were obtained when production cost is the same in each period.

In our first set of experiments, we study the impact of the percentage of demand from the first class. In these cases, the expected demand from the first class customers takes on the values of 10, 30, 50, 70, and 90, and the second class demand is $100 - E(D^1)$. The prices are constant over the set of experiments but may vary by period. The average ratio of p^1/p^2 is fixed at 1.2 for these experiments. Please refer to Table 3.1 in Appendix 3.9 for the data used in this experiment.

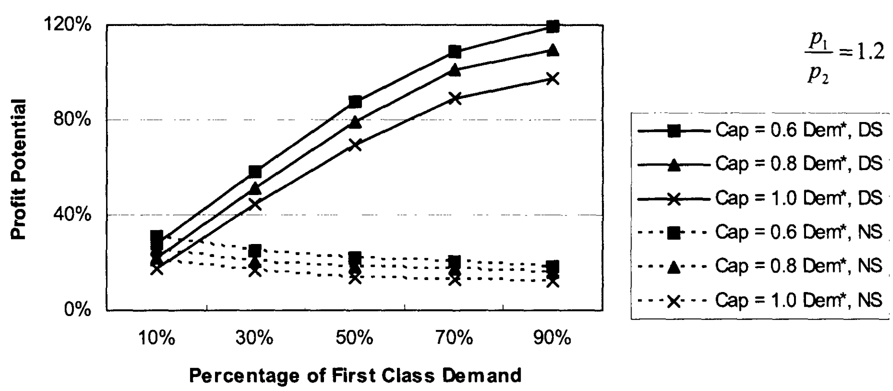


Figure 3-1: Impact of Demand Ratio

In Figure 3-1, we find that both the Differentiation and the Non-differentiation Strategies result in a higher profit than the traditional policy. The relative performance of each (S, R, B) policy increases as the production capacity decreases. This is because the (S, R, B) policies can manage the inventory more effectively than the traditional policy, and their impacts become more prominent as the production capacity becomes a scarce resource.

The Differentiation Strategy outperforms the Nondifferentiation Strategy in most situations except when the 1^{st} class has a very low proportion among the customers (Figure 3-1). In this situation, under the Differentiation Strategy, the extra revenue gained from the 1^{st} class is not high enough to cover the additional inventory holding cost and backlogging cost caused by the delayed fulfillments of the 2^{nd} class demand. In general, the profit improvement under the Differentiation Strategy is much higher than that under the Nondifferentiation Strategy, therefore we can see that the market segmentation factor brings much higher profit improvement than the shifting inventory factor.

For a given capacity level, the profit potential of the Differentiation Strategy increases dramatically as the proportion of the 1^{st} class customers increases, because the 1^{st} class demand can bring high profit to the manufacturer. In contrast, the profit potential of the Nondifferentiation Strategy decreases as the proportion of 1^{st} class customers increases, because more demand will be lost when the inventory is not available.

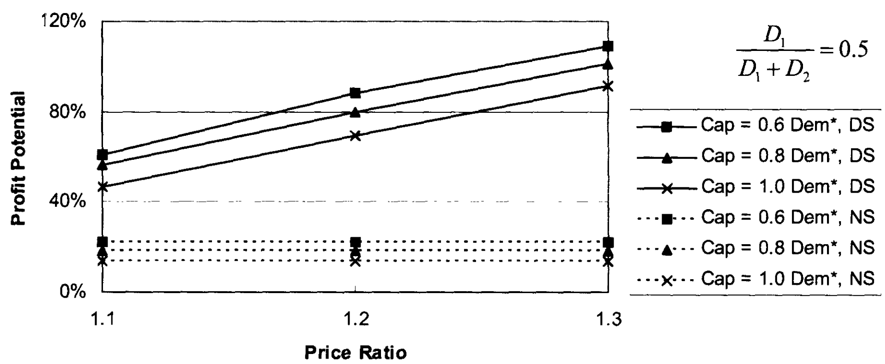


Figure 3-2: Impact of Price Ratio

In the second set of experiments, we study the impact of price difference between the two classes. In this set of experiments, $E(p^2)$ is fixed over the instances, and p^1 is set according to $E(p^1)/E(p^2) = 1.1$, 1.2, and 1.3, where the E represents the average price over the horizon. The percentage of demand from the first class is fixed at 50%.

In Figure 3-2, we see that the relative performance of the Differentiation Strategy increases as the price ratio increases. The relative performance of the Nondifferentiation Strategy stays unchanged since p_t^1 affects neither the Nondifferentiation Strategy nor the traditional policy.

In the numerical analysis, we also find that the heuristic policy under the Differentiation Strategy performs close to the optimal policy. Among all the experiments we have done, Condition 3.2 is satisfied in 93%, in which case the heuristic is the optimal policy. Condition 3.2 may not be satisfied when the percentage of demand from the first class is very low, i.e., $\frac{D_1}{D_1+D_2} \leq 30\%$.

3.6 Concluding Remarks

In this chapter, we study a single-product, two-class-customer problem with both lost sales and partial backorders under two strategies: under the Non-differentiation Strategy, the manufacturer does not differentiate customers and serves the customers as a single class; under the Differentiation Strategy, the manufacturer provides different qualities of service and charge different prices to different classes of customers.

For the Non-differentiation Strategy, we characterize the structure of the optimal policy with three state-independent parameters: the base-stock level, S, the reserve-up-to level, R, and the backlog-up-to level, B.

Computational analysis shows that the Differentiation Strategy can increase the manufacturer's profit significantly, especially when the production capacity is tight, when the percentage of waiting customers is high, and when price difference between the two classes is high.

3.7 Appendix A

Proof of Lemma 3.1

Proof. By contradiction, assume that there is an optimal policy with $R_t \cdot B_t > 0$ for some period t. Let $\bar{R}_t = R_t - 1$ and $\bar{B}_t$ be the alternative policy, and let V_t and $\bar{V}_t$ be the expected profit starting from period t under the two policies respectively. We compare the two policies in the following three cases:

- Case 1: $D_t \leq Y_t - R_t$, hence $D_t < Y_t - \bar{R}_t$.

$$\begin{aligned} V_t &= p_t^2 D_t - h_t(Y_t - D_t) + J_{t+1}(Y_t - D_t) \\ \bar{V}_t &= p_t^2 D_t - h_t(Y_t - D_t) + J_{t+1}(Y_t - D_t) = V_t \end{aligned}$$

- Case 2: $Y_t - R_t + B_t/\alpha_t > D_t > Y_t - R_t$, hence $Y_t - \bar{R}_t + \bar{B}_t/\alpha_t \geq D_t \geq Y_t - R_t$.

$$\begin{aligned} V_t &= p_t^2(Y_t - R_t + \lfloor \alpha_t(D_t - Y_t + R_t) \rfloor) - h_t R_t - b_t^2 \lfloor \alpha_t(D_t - Y_t + R_t) \rfloor \\ &\quad -r_t^1 \lceil (1-\alpha_t)(D_t - Y_t + R_t) \rceil + J_{t+1}(R_t - \lfloor \alpha_t(D_t - Y_t + R_t) \rfloor) \\ \bar{V}_t &= p_t^2(Y_t - R_t + 1 + \lfloor \alpha_t(D_t - Y_t + R_t - 1) \rfloor) - h_t(R_t - 1) - b_t^2 \lfloor \alpha_t(D_t - Y_t + R_t - 1) \rfloor \\ &\quad -r_t^1 \lceil (1-\alpha_t)(D_t - Y_t + R_t - 1) \rceil + J_{t+1}(R_t - 1 - \lfloor \alpha_t(D_t - Y_t + R_t - 1) \rfloor) \end{aligned}$$

If $\lfloor \alpha_t(D_t - Y_t + R_t - 1) \rfloor = \lfloor \alpha_t(D_t - Y_t + R_t) \rfloor$, then $\lceil (1-\alpha_t)(D_t - Y_t + R_t - 1) \rceil = \lceil (1-\alpha_t)(D_t - Y_t + R_t) \rceil - 1$. We have,

$$\bar{V}_t = V_t + p_t^2 + h_t + r_t^1 - J_{t+1}(R_t - \lfloor \alpha_t(D_t - Y_t + R_t) \rfloor) + J_{t+1}(R_t - 1 - \lfloor \alpha_t(D_t - Y_t + R_t) \rfloor)$$

Since $D_t < Y_t - R_t + B_t/\alpha_t$, a new demand from class 2 will be accepted, which means $p_t^2 + h_t + r_t^2 + J_{t+1}(R_t - 1 - \lfloor \alpha_t(D_t - Y_t + R_t) \rfloor) \geq J_{t+1}(R_t - \lfloor \alpha_t(D_t - Y_t + R_t) \rfloor)$. Thus $\bar{V}_t \geq V_t$.

Otherwise, $\lfloor \alpha_t(D_t - Y_t + R_t - 1) \rfloor = \lfloor \alpha_t(D_t - Y_t + R_t) \rfloor$-1, then $\lceil (1-\alpha_t)(D_t - Y_t + R_t - 1) \rceil = \lceil (1-\alpha_t)(D_t - Y_t + R_t) \rceil$. We have,

$$\bar{V}_t = V_t + h_t + b_t^2 \geq V_t$$

- Case 3: $D_t \geq Y_t - R_t + B_t/\alpha_t$, hence $D_t > Y_t - \bar{R}_t + \bar{B}_t/\alpha_t$.

$$\begin{aligned} V_t &= p_t^2(Y_t - R_t + B_t) - h_t R_t - b_t^2 B_t - r_t^1(1-\alpha_t)B_t/\alpha_t - r_t(D_t - Y_t + R_t - B_t/\alpha_t) \\ &\quad +J_{t+1}(R_t - B_t) \\ \bar{V}_t &= p_t^2(Y_t - R_t + 1 + B_t - 1) - h_t(R_t - 1) - b_t^2(B_t - 1) - r_t^1(1-\alpha_t)(B_t - 1)/\alpha_t \\ &\quad -r_t(D_t - Y_t + R_t - 1 - (B_t - 1)/\alpha_t) + J_{t+1}(R_t - B_t) \\ &= V_t + h_t + b_t^2 + (1-\alpha_t)(r_t^1 - r_t^2) \\ &\geq V_t \end{aligned}$$

The expected profit under the alternative policy is always grater or equal to that under the current policy, which incurs a contradiction.

The following lemma is used in the proof for Theorem 3.1.

Lemma 3.6 *Given $g(x,y)$ is jointly concave in x and y, $G(x)=\max_y g(x,y)$ is a concave function for x.*

Proof. For any $x_1, x_2 \in R$, let $y_1 = \arg\max\{y|g(x_1,y)\}$, $y_2 = \arg\max\{y|g(x_2,y)\}$. For any $\lambda \in [0,1]$, let $x_\lambda = \lambda x_1 + (1-\lambda)x_2$, $y_\lambda = \lambda y_1 + (1-\lambda)y_2$. We have $G(x_\lambda) = \max_y g(x_\lambda, y) \geq g(x_\lambda, y_\lambda) \geq \lambda g(x_1,y_1) + (1-\lambda)g(x_2,y_2) = \lambda G(x_1) + (1-\lambda)G(x_2)$. ■

Proof of Concavity for Nondifferentiation Strategy (Theorem 3.1)

Proof. Let $j_t(I_t,Y_t) = -c_t(Y_t - I_t) + G_t(Y_t)$, so $J_t(I_t) = \max_{Y_t : I_t \leq Y_t \leq I_t + q_t} j_t(I_t,Y_t)$. We prove by induction.

1. For period T, we have $J_{T+1}(I_T) = vI_T$, $B_T = 0$, $R_T = 0$, and therefore $G_T^R(Y_T) = G_T^B(Y_T) = G_T(Y_T)$.

   $G_T(Y_T) = \int\{p_T^2 \min(D_T,Y_T) - r_T(D_T - Y_T)^+ - h_T(Y_T - D_T)^+ + v(Y_T - D_T)^+\}d\Phi_T$.

   $G_T'(Y_T) = \int_{Y_T}^{\infty}(p_T^2 + r_T)d\Phi_T + \int_0^{Y_T}(v - h_T)d\Phi_T$. Thus $G_T(Y_T)$ is concave in Y_T since $G_T'(Y_T)$ is non-increasing in Y_T:

   $G_T''(Y_T) = (v - h_T - p_T^2 - r_T)\phi_T(Y_T) \leq 0$.

2. Given t, $T = t, ..., T$, assume that $G_{t+1}(Y_T)$ is concave in Y_T, then we can prove that $j_{t+1}(I_{t+1},Y_{t+1})$ is jointly concave in I_{t+1} and Y_{t+1} by the following. For any $(I_1,Y_1),(I_2,Y_2)$, let $I_\lambda = \lambda I_1 + (1-\lambda)I_2$, $Y_\lambda = \lambda Y_1 + (1-\lambda)Y_2$. Then,

$$\begin{aligned} j_{t+1}(I_\lambda,Y_\lambda) &= -c_{t+1}(Y_\lambda - I_\lambda) + G_{t+1}(Y_\lambda) \\ &= -c_{t+1}(\lambda Y_1 + (1-\lambda)Y_2 - \lambda I_1 - (1-\lambda)I_2) + G_{t+1}(\lambda Y_1 + (1-\lambda)Y_2) \\ &\geq -\lambda c_{t+1}(Y_1 - I_1) - (1-\lambda)c_{t+1}(Y_2 - I_2) + \lambda G_{t+1}(Y_1) + (1-\lambda)G_{t+1}(Y_2) \\ &= \lambda j_{t+1}(I_1,Y_1) + (1-\lambda)j_{t+1}(I_2,Y_2). \end{aligned}$$

   So by Lemma 3, $J_{t+1}(I_t)$ is concave in I_t, and as a result $J_{t+1}'(I_t)$ is non-increasing in I_t.

3. Next let us prove that $g_t^R(Y_t,R_t)$ is quasi-concave in R_t.

$$g_t^{'R}(R_t) = \begin{cases} 0 & if\ 0 \le Y_t < R_t \\ (-p_t^2 - r_t - h_t + J_{t+1}^{'}(R_t))(1 - \Phi_t(Y_t - R_t)) & if\ Y_t \ge R_t. \end{cases}$$

Let us define R_t^* as

$$R_t^* = \begin{cases} \max\{I : J_{t+1}^{'}(I) \le p_t^2 + r_t + h_t\} & if\ (p_t^2 + r_t + h_t) < J_{t+1}^{'}(0) \\ 0 & otherwise. \end{cases}$$

Thus we have $g_t^{'R}(R_t) \ge 0$, when $0 \le R_t \le R_t^*$, and $g_t^{'R}(R_t) \le 0$, when $R_t > R_t^*$, so, $g_t^R(Y_t, R_t)$ is quasi-concave with respect to R_t. R_t^* is the unique unconstrained optimizer of $g_t^R(Y_t, R_t)$ and it is independent of inventory level Y_t. $R_t^c = \min(R_t^*, Y_t)$ maximizes $g_t^R(Y_t, R_t)$, for $0 \le R_t \le (Y_t)^+$.

4. Next let us prove that $g_t^B(Y_t, B_t)$ is quasi-concave in B_t. Take the derivative,

$$g_t^{'B}(B_t) = \int_{Y_t + B_t/\alpha}^{\infty} [p_t^2 - b_t^2 + r_t^2 - J_{t+1}'(-B_t)] d\Phi_t.$$

Let us define B_t^* as

$$B_t^* = \begin{cases} \max\{I : p_t^2 + r_t^2 - b_t^2 \ge J_{t+1}^{'}(-I)\} & if\ (p_t^2 + r_t^2 - b_t^2) > J_{t+1}^{'}(0) \\ 0 & otherwise. \end{cases}$$

Thus we have $g_t^{'B}(B_t) \ge 0$, when $0 \le B_t \le B_t^*$, and $g_t^{'B}(B_t) \le 0$, when $B_t > B_t^*$, so, $g_t^B(Y_t, B_t)$ is quasi-concave with respect to B_t. B_t^* is the unique unconstrained optimizer of $g_t^B(Y_t, B_t)$ and it is independent of inventory level Y_t. $B_t^c = \min(B_t^*, q_{t+1})$ maximizes $g_t^B(Y_t, B_t)$, for $0 \le B_t \le q_{t+1}$.

5. Let us prove the concavity of $G_t^R(Y_t)$ with respect to Y_t.

   – First let us examine the derivative of $G_t^R(Y_t)$ with respect to Y_t.

$$\begin{aligned} G_t^R(Y_t) &= g_t^R(Y_t, R_t^c) \\ &= \int \{p_t^2 \min(D_t, Y_t - R_t^c) - h_t \max(R_t^c, Y_t - D_t) - r_t(D_t - Y_T + R_t^c)^+ \\ &\quad + J_{t+1}(\max(R_t^c, Y_t - D_t))\} d\Phi_t; \end{aligned}$$

$$G_t^{'R}(Y_t) = \begin{cases} J_{t+1}'(Y_t) - h_t & if\ 0 \le Y_t \le R_t^* \\ \int_{Y_t - R_t^*}^{\infty} (p_t^2 + r_t) d\Phi_t + \int_0^{Y_t - R_t^*} (J_{t+1}'(Y_t - D_t) - h_t) d\Phi_t & if\ Y_t > R_t^*. \end{cases}$$

- We consider $G_t''^R(Y_t)$ in three cases:

  (a)Case 1: $Y_t < R_t^*$:

  $G_t''^R(Y_t) = J_{t+1}''(Y_t) \leq 0$, due to the concavity of J_{t+1}.

  (b)Case 2: $Y_t > R_t^*$:

  $G_t''^R(Y_t) = \int_0^{Y_t - R_t^*} J_{t+1}''(Y_t - D_t)d\Phi_t + (J_{t+1}'(R_t^*) - p_t^2 - r_t - h_t)\phi_t(Y_t - R_t^*) \leq 0$

  due to the choice of R_t^* and the concavity of J_{t+1}.

  (c)Case 3: $Y_t = R_t^*$:

  $G_t'R(Y_t+) - G_t'R(Y_t-) = p_t^2 + r_t + h_t - J_{t+1}'(R_t^*) \leq 0$

  due to the choice of R_t^* and the concavity of J_{t+1}.

- Since $G_t''^R(Y_t) \leq 0$ for all Y_t, $G_t^R(Y_t)$ is concave in Y_t.

6. Let us prove the concavity of $G_t^B(Y_t)$ with respect to Y_t.

$$
\begin{aligned}
G_t^B(Y_t) &= g_t^B(Y_t, B_t^c) \\
&= \int \{p_t^2 \min(D_t, Y_t + \min(B_t^c, \alpha(D_t - Y_t)^+))^+ - h_t(Y_t - D_t)^+ \\
&\quad - r_t^1(1-\alpha)\min(B_t^c/\alpha, (D_t - Y_t)^+) - r_t(D_t - Y_t - B_t^c/\alpha)^+ \\
&\quad + J_{t+1}((Y_t - D_t)^+ - \min(\alpha(D_t - Y_t)^+, B_t^c)) - b_t^2 \min(B_t^c, \alpha(D_t - Y_t)^+)\} d\Phi_t;
\end{aligned}
$$

$$
\begin{aligned}
G_t'^B(Y_t) &= \int_{Y_t}^{Y_t + B_t/\alpha} [(1-\alpha)(p_t^2 + r_t^1) + \alpha(b_t^2 + J_{t+1}'(\alpha(Y_t - D_t)))] d\Phi_t \\
&\quad + \int_0^{Y_t} (-h_t + J_{t+1}'(Y_t - D_t)) d\Phi_t + \int_{Y_t + B/\alpha}^{\infty} (p_t^2 + r_t) d\Phi_t;
\end{aligned}
$$

$$
\begin{aligned}
G_t''^B(Y_t) &= -\alpha(p_t^2 + r_t^2 - b_t^2 - J_{t+1}'(-B_t))\phi_t(Y_t + B_t/\alpha) \\
&\quad + [\alpha(p_t^2 + r_t^2 - b_2 - J_{t+1}'(0)) - (p_t^2 + r_t^1 + h_t - J_{t+1}'(0))]\phi_t(Y_t) \\
&\quad + \int_0^{Y_t} J_{t+1}''(Y_t - D_t) d\Phi_t + \int_Y Y_t^{Y_t + B_t/\alpha} \alpha^2 J_{t+1}''(\alpha(Y_t - D_t)) d\Phi_t.
\end{aligned}
$$

The first term in $G_t''^B(Y_t)$ is negative due to the choice of B_t^*. The third and the fourth terms in $G_t''^B(Y_t)$ are negative due to the concavity of $J_{t+1}(Y_t)$. We have $G_t''^B(Y_t) \leq 0$ and therefore, $G_t^B(Y_t, B_t)$ is concave in Y_t.

7. Let us prove the concavity of $G_t(Y_t)$.

   In each period, we must be in one of the following cases; which are independent of the Y_t values:

   - If $p_t^2 + r_t^2 - b_t^2 \leq J_{t+1}'(0) \leq p_t^2 + r_t + h_t$, we have $R_t^* = B_t^* = 0$, therefore $R_t^c = B_t^c = 0$, so we have $G_t(Y_t) = G_t^R(Y_t) = G_t^B(Y_t)$, which can be seen from the formulations directly.

- If $J'_{t+1}(0) > p_t^2 + r_t + h_t$, we have $R_t^* \geq 0$ and $B_t^* = 0$, therefore $R_t^c \geq 0$ and $B_t^c = 0$, so we have $G_t(Y_t) = G_t^R(Y_t) \geq G_t^B(Y_t)$.
- If $J'_{t+1}(0) < p_t^2 + r_t^2 - b_t^2 <$, we have $B_t^* \geq 0$ and $R_t^* = 0$, therefore $B_t^c \geq 0$ and $R_t^c = 0$, so we have $G_t(Y_t) = G_t^B(Y_t) \geq G_t^R(Y_t)$.

We see that in each period, $G_t(Y_t)$ reduces to some function that is proved to be concave. Therefore $G_t(Y_t)$ is concave. ■

3.8 Appendix B

Proof of Lemma 3.2

Proof. By contradiction, assume that there is an optimal policy with $R_t \cdot B_t > 0$ for some period t. Let $\bar{R}_t = R_t - 1$ and $\bar{B}_t = B_t - 1$ be the alternative policy, and let V_t and $\bar{V}_t$ be the expected profit starting from period t under the two policies respectively. We compare the two policies in the following four cases:

- Case 1: $D_t^1 > Y_t - R_t$ and $D_t^2 \geq B_t$, hence $D_t^1 \geq Y_t - \bar{R}_t$ and $D_t^2 > \bar{B}_t$.

$$\begin{aligned} V_t &= p_t^1(Y_t - R_t) + (p_t^2 - b_t^2)B_t - h_t R_t - r_t^1(D_t^1 - Y_t + R_t) - r_t^2(D_t^2 - B_t) + J_{t+1}(R_t - B_t) \\ \bar{V}_t &= p_t^1(Y_t - \bar{R}_t) + (p_t^2 - b_t^2)\bar{B}_t - h_t \bar{R}_t - r_t^1(D_t^1 - Y_t + \bar{R}_t) - r_t^2(D_t^2 - \bar{B}_t) + J_{t+1}(\bar{R}_t - \bar{B}_t) \\ &= V_t + p_t^1 + r_t^1 + h_t - (p_t^2 - b_t^2 + r_t^2) \\ &\geq V_t \end{aligned}$$

- Case 2: $D_t^1 > Y_t - R_t$ and $D_t^2 < B_t$, hence $D_t^1 \geq Y_t - \bar{R}_t$ and $D_t^2 \leq \bar{B}_t$.

$$\begin{aligned} V_t &= p_t^1(Y_t - R_t) + (p_t^2 - b_t^2)D_t^2 - h_t R_t - r_t^1(D_t^1 - Y_t + R_t) + J_{t+1}(R_t - D_t^2) \\ \bar{V}_t &= p_t^1(Y_t - \bar{R}_t) + (p_t^2 - b_t^2)D_t^2 - h_t \bar{R}_t - r_t^1(D_t^1 - Y_t + \bar{R}_t) + J_{t+1}(\bar{R}_t - D_t^2) \\ &= V_t + p_t^1 + r_t^1 + h_t - J_{t+1}(R_t - D_t^2) + J_{t+1}(R_t - 1 - D_t^2) \end{aligned}$$

  Since $D_t^2 < B_t$, one more unit of class 2 demand will be accepted if it arrives, which means $p_t^2 - b_t^2 + J_{t+1}(R_t - 1 - D_t^2) \geq J_{t+1}(R_t - D_t^2) - r_t^2$. Since $p_t^1 + r_t^1 + h_t \geq p_t^2 - b_t^2 + r_t^2$, $\bar{V}_t \geq V_t$.

- Case 3: $D_t^1 \leq Y_t - R_t$ and $D_t^2 \geq B_t$, hence $D_t^1 < Y_t - \bar{R}_t$ and $D_t^2 > \bar{B}_t$.

$$\begin{aligned} V_t &= p_t^1 D_t^1 + (p_t^2 - b_t^2)B_t - h_t(Y_t - D_t^1) - r_t^2(D_t^2 - B_t) + J_{t+1}(Y_t - D_t^1 - B_t) \\ \bar{V}_t &= p_t^1 D_t^1 + (p_t^2 - b_t^2)\bar{B}_t - h_t(Y_t - D_t^1) - r_t^2(D_t^2 - \bar{B}_t) + J_{t+1}(Y_t - D_t^1 - \bar{B}_t) \\ &= V_t - (p_t^2 - b_t^2 + r_t^2) - J_{t+1}(Y_t - D_t^1 - B_t) + J_{t+1}(Y_t - D_t^1 - B_t + 1) \end{aligned}$$

  Since $D_t^2 \geq B_t$, class 2 demands beyond B_t will be rejected, which means $p_t^2 - b_t^2 + J_{t+1}(Y_t - D_t^1 - B_t + 1) \geq J_{t+1}(Y_t - D_t^1 - B_t) - r_t^2$, thus $\bar{V}_t \geq V_t$.

- Case 4: $D_t^1 \leq Y_t - R_t$ and $D_t^2 < B_t$, hence $D_t^1 < Y_t - \bar{R}_t$ and $D_t^2 \leq \bar{B}_t$.

$$\begin{aligned} V_t &= p_t^1 D_t^1 + (p_t^2 - b_t^2) D_t^2 - h_t(Y_t - D_t^1) + J_{t+1}(Y_t - D_t^1 - D_t^2) \\ \bar{V}_t &= p_t^1 D_t^1 + (p_t^2 - b_t^2) D_t^2 - h_t(Y_t - D_t^1) + J_{t+1}(Y_t - D_t^1 - D_t^2) \\ &= V_t \end{aligned}$$

The expected profit under the alternative policy is always greater or equal to that under the current policy, which incurs a contradiction. ■

Proof of Quasi-Concavity in R_t (Lemma 3.3)

Proof. The first order derivative of $g_t^R(Y_t, R_t)$ with respect to R_t is,

$$g_R^{'} = \frac{\partial g_t^R}{\partial R} = \begin{cases} 0 & if\ 0 \leq Y_t < R_t \\ (-p_t^1 - r_t^1 - h_t + J_{t+1}^{'}(R_t))(1 - \Phi_t^1(Y_t - R_t)) & if\ Y_t \geq R_t. \end{cases}$$

Let us define R_t^* as

$$R_t^* = \begin{cases} \max\{I : p_t^1 + r_t^1 + h_t < J_{t+1}^{'}(I)\} & if\ (p_t^1 + r_t^1 + h_t) < J_{t+1}^{'}(0) \\ 0 & otherwise. \end{cases} \tag{3.16}$$

Thus we have $g_t^{'R}(R_t) \geq 0$, when $0 \leq R_t \leq R_t^*$, and $g_t^{'R}(R_t) \leq 0$, when $R_t > R_t^*$, so, $g_t^R(Y_t, R_t)$ is quasi-concave with respect to R_t. R_t^* is the unique unconstrained optimizer of $g_t^R(Y_t, R_t)$ and it is independent of inventory level Y_t. $R_t^c = \min(R_t^*, Y_t)$ maximizes $g_t^R(Y_t, R_t)$, for $0 \leq R_t \leq Y_t$. ■

Before proving Lemma 3.4, let us first prove the following lemma.

Lemma 3.7 *For any Y_t, if $g_B^{'}(B_t, Y_t) \leq 0$, then $g_B^{'}(B_t + \epsilon, Y_t) \leq 0$, where ϵ is a positive infinite small number.*

Proof. The first and second order derivatives of $g_t^B(Y_t, B_t)$ with respect to B_t are,

$$\begin{aligned} g_B^{'} = \frac{\partial g_t^B}{\partial B} &= \int_0^{Y_t} \int_{B_t}^{\infty} [p_t^2 - \ell_t^2 + \beta_t^2 - J_{t+1}^{'}(Y_t - k_1 - B_t)] d\Phi_t^2(k_2) d\Phi_t^1(k_1) \\ &+ \int_{Y_t}^{\infty} \int_{B_t}^{\infty} [p_t^2 - \ell_t^2 + \beta_t^2 - J_{t+1}^{'}(-B_t)] d\Phi_t^2(k_2) d\Phi_t^1(k_1). \end{aligned} \tag{3.17}$$

$$
\begin{aligned}
g''_{BB} = \frac{\partial g'_B}{\partial B} &= \int_0^{Y_t} \int_{B_t}^{\infty} [J''_{t+1}(Y_t - k_1 - B_t)] d\mathbf{\Phi}_t^2(k_2) d\mathbf{\Phi}_t^1(k_1) \\
&+ \int_{Y_t}^{\infty} \int_{B_t}^{\infty} [J''_{t+1}(-B_t)] d\mathbf{\Phi}_t^2(k_2) d\mathbf{\Phi}_t^1(k_1) \\
&+ \int_0^{Y_t} [p_t^2 - \ell_t^2 + \beta_t^2 - J'_{t+1}(Y_t - k_1 - B_t)] d\mathbf{\Phi}_t^1(k_1) \\
&+ \int_{Y_t}^{\infty} [p_t^2 - \ell_t^2 + \beta_t^2 - J'_{t+1}(-B_t)] d\mathbf{\Phi}_t^1(k_1) \qquad (3.18)
\end{aligned}
$$

For simplification, let us denote,

$$
\begin{aligned}
V_t(B_t, Y_t) &= \int_0^{Y_t} \int_{B_t}^{\infty} [J''_{t+1}(Y_t - k_1 - B_t)] d\mathbf{\Phi}_t^2(k_2) d\mathbf{\Phi}_t^1(k_1) + \int_{Y_t}^{\infty} \int_{B_t}^{\infty} [J''_{t+1}(-B_t)] d\mathbf{\Phi}_t^2(k_2) d\mathbf{\Phi}_t^1(k_1) \\
W_t(B_t, Y_t) &= \int_0^{Y_t} [p_t^2 - \ell_t^2 + \beta_t^2 - J'_{t+1}(Y_t - k_1 - B_t)] d\mathbf{\Phi}_t^1(k_1) + \int_{Y_t}^{\infty} [p_t^2 - \ell_t^2 + \beta_t^2 - J'_{t+1}(-B_t)] d\mathbf{\Phi}_t^1(k_1).
\end{aligned}
$$

Then we have,

$$
\begin{aligned}
g'_B(B_t, Y_t) &= W_t(B_t, Y_t)[1 - \mathbf{\Phi}_t^2(B_t)] \qquad (3.19) \\
g''_{BB}(B_t, Y_t) &= V_t(B_t, Y_t) - W_t(B_t, Y_t) \qquad (3.20)
\end{aligned}
$$

By contradiction, assume exists a B_t, such that $g'_B(B_t, Y_t) \leq 0$ and $g'_B(B_t + \epsilon, Y_t) > 0$. By (3.19), we have,

$$
W_t(B_t + \epsilon, Y_t) = g'_B(B_t + \epsilon, Y_t)/[1 - \mathbf{\Phi}_t^2(B_t + \epsilon)] > 0.
$$

Since J_{t+1} is concave, we have $V_t(B_t + \epsilon, Y_t) \leq 0$, and by (3.20),

$$
g''_{BB}(B_t + \epsilon, Y_t) = V_t(B_t + \epsilon, Y_t) - W_t(B_t + \epsilon, Y_t) < 0.
$$

By definition,

$$
g''_{BB}(B_t + \epsilon, Y_t) = \frac{g'_B(B_t + \epsilon, Y_t) - g'_B(B_t, Y_t)}{\epsilon},
$$

We have

$$
g'_B(B_t, Y_t) = g'_B(B_t + \epsilon, Y_t) - \epsilon g''_{BB}(B_t + \epsilon, Y_t) > 0.
$$

Contradiction achieved. ■

Proof of Quasi-Concavity in B_t (Lemma 3.4)

Proof. Lemma 3.7 implies that for a given value of Y_t, once $g_t^B(B_t, Y_t)$ becomes non-increasing in B_t, it will keep non-increasing as B_t further increases. In other words, $g_t^B(B_t, Y_t)$ satisfies one out of the following three conditions: (1) non-decreasing in B_t; (2)

non-increasing in B_t; (3) or there exists $B_t^*(Y_t)$ such that it is non-decreasing for $B_t < B_t^*(Y_t)$ and non-increasing for $B_t > B_t^*(Y_t)$. Therefore $g_t^B(B_t, Y_t)$ is quasi-concave in B_t. ■

Proof of Property 3.1

Proof. The proof of Property 3.1 is straightforward by noticing that in (3.16), R_t^* is independent of Y_t. ■

Proof of Property 3.2

Proof.

- Let us first examine the derivative of $G_t^R(Y_t)$ with respect to Y_t.

$$\begin{aligned} G_t^R(Y_t) = g_t^R(Y_t, R_t^c) = & \int \{p_t^1 \min(D_t^1, Y_t - R_t^c) - h_t \max(R_t^c, Y_t - D_t^1) \\ & -r_t^1(D_t^1 - Y_t + R_t^c)^+ + J_{t+1}(\max(R_t^c, Y_t - D_t^1))\} d\Phi_t^1. \end{aligned}$$

$$G_t'^R(Y_t) = \begin{cases} J_{t+1}'(Y_t) - h_t & if\ 0 \le Y_t \le R_t^* \\ \int_{Y_t - R_t^*}^{\infty} (p_t^1 + r_t^1) d\Phi_t^1 + \int_0^{Y_t - R_t^*} (J_{t+1}'(Y_t - D_t^1) - h_t) d\Phi_t^1 & if\ Y_t > R_t^*. \end{cases}$$

- We consider $G_t''^R(Y_t)$ in three cases:

  (a)Case 1: $Y_t < R_t^*$:

  $G_t''^R(Y_t) = J_{t+1}''(Y_t) \le 0$, due to the concavity of J_{t+1}.

  (b)Case 2: $Y_t > R_t^*$:

  $G_t''^R(Y_t) = \int_0^{Y_t - R_t^*} J_{t+1}''(Y_t - D_t^1) d\Phi_t^1 + (J_{t+1}'(R_t^*) - p_t^1 - r_t^1 - h_t)\phi_t^1(Y_t - R_t^*) \le 0$

  due to the choice of R_t^* and the concavity of J_{t+1}.

  (c)Case 3: $Y_t = R_t^*$:

  $G_t''^R(Y_t) = p_t^1 + r_t^1 + h_t - J_{t+1}'(R_t^*) \le 0$

  due to the choice of R_t^* and the concavity of J_{t+1}.

- Since $G_t''^R(Y_t) \le 0$ for all Y_t, $G_t^R(Y_t)$ is concave in Y_t.

■

Proof of Property 3.3

Proof. The proof of Property 3.3 is straightforward by plugging (3.13) into (3.17). ■

Proof of Property 3.5

Proof. To prove the joint quasi-concavity, we study the determinants of the bordered Hessian:

$$D_1 = \begin{vmatrix} 0 & g_B' \\ g_B' & g_{BB}'' \end{vmatrix}$$

$$D_2 = \begin{vmatrix} 0 & g_B' & g_Y' - c_t \\ g_B' & g_{BB}'' & g_{BY}'' \\ g_Y' - c_t & g_{BY}'' & g_{YY}'' \end{vmatrix}$$

It is easy to show $|D_1| = -(g_B')^2 \leq 0$, so we only need to show that,

$$|D_2| = 2g_B'(g_Y' - c_t)g_{BY}'' - g_B'g_B'g_{YY}'' - (g_Y' - c_t)(g_Y' - c_t)g_{BB}'' \geq 0$$

. Denote $N = g_{BY}'' = \frac{\partial g_B'}{\partial Y} = \int_0^{Y_t}\int_{B_t}^{\infty}[J_{t+1}''(Y_t - k_1 - B_t)]d\mathbf{\Phi}_t^2(k_2)d\mathbf{\Phi}_t^1(k_1)$, so $N \geq 0$. Notice that N also appears in $-g_{YY}''$ and $-g_{BB}''$. So can take the following term out from $|D_2|$,

$$M = 2g_B'(g_Y' - c_t)N + g_B'g_B'N + (g_Y' - c_t)(g_Y' - c_t)N = N(g_B' + g_Y' - c_t)^2 \geq 0$$

So if we can prove $|D_2| - M \geq 0$, then it will be sufficient to show that $|D_2| \geq 0$.

$$\begin{aligned}
|D_2| - M \quad = \quad & -g_B'g_B'\{\int_0^{Y_t}\int_0^{B_t}[J_{t+1}''(Y_t - k_1 - k_2)]d\mathbf{\Phi}_t^2(k_2)d\mathbf{\Phi}_t^1(k_1) \\
& + \int_0^{B_t}[-p_t^1 - \beta_t^1 - h_t + J_{t+1}'(-k_2)]d\mathbf{\Phi}_t^2(k_2) \\
& + \int_{B_t}^{\infty}[-p_t^1 - \beta_t^1 - h_t + J_{t+1}'(-B_t)]d\mathbf{\Phi}_t^2(k_2)\} \\
& -(g_Y' - c_t)(g_Y' - c_t)\{\int_{Y_t}^{\infty}\int_{B_t}^{\infty}[J_{t+1}''(-B_t)]d\mathbf{\Phi}_t^2(k_2)d\mathbf{\Phi}_t^1(k_1) \\
& - \int_0^{Y_t}[p_t^2 - \ell_t^2 + \beta_t^2 - J_{t+1}'(Y_t - k_1 - B_t)]d\mathbf{\Phi}_t^1(k_1) \\
& - \int_{Y_t}^{\infty}[p_t^2 - \ell_t^2 + \beta_t^2 - J_{t+1}'(-B_t)]d\mathbf{\Phi}_t^1(k_1)\}
\end{aligned}$$

The first and the fourth terms are negative due to the concavity of J_{t+1}. The second and the third terms are negative under Condition 3.2. The summation of the last two terms equal to $-g_B'/[1 - \Phi_2(B_t)]$, so it is negative when $B_t \leq B_t^*(Y_t)$. So $|D_2| - M \geq 0$, and it completes the proof for Lemma 3.5. ∎

Proof of Property 3.4

Proof. Let us prove the concavity of $G_t^B(Y_t)$ with respect to Y_t.

$$g_t^B(Y_t, B_t^c) \quad = \quad \int\int\{p_t^1\min(D_t^1, Y_t) - h_t(Y_t - D_t^1)^+ - r_t^1(D_t^1 - Y_t)^+ - r_t^2(D_t^2 - B_t^c)^+$$

$$
\begin{aligned}
& +(p_t^2 - b_t^2)\min(D_t^2, B_t^c) + J_{t+1}((Y_t - D_t^1)^+ - \min(D_t^2, B_t^c))\}d\Phi_t^1 d\Phi_t^2 \\
G_t'^B(Y_t) \quad = \quad & \int_0^{Y_t} d\Phi_t^1\{\int_0^B (J'_{t+1}(Y_t - D_t^1 - D_t^2) - h_t)d\Phi_t^2 \\
& + \int_B^\infty (J'_{t+1}(Y_t - D_t^1 - B_t) - h_t)d\Phi_t^2\} + \int_{Y_t}^\infty (p_t^1 + r_t^1)d\Phi_t^1 \\
G_t''^B(Y_t) \quad = \quad & \int_0^{Y_t} d\Phi_t^1\{\int_0^{B_t} J''_{t+1}(Y_t - D_t^1 - D_t^2)d\Phi_t^2 + \int_{B_t}^\infty J''_{t+1}(Y_t - D_t^1 - B_t)d\Phi_t^2\} \\
& +\phi_t^1(Y_t)\{\int_0^{B_t} (J'_{t+1}(-D_t^2) - h_t - r_t^1 - p_t^1)d\Phi_t^2\} \\
& +\phi_t^1(Y_t)\{\int_{B_t}^\infty (J'_{t+1}(-B_t) - h_t - r_t^1 - p_t^1)d\Phi_t^2\}
\end{aligned}
$$

If Condition 3.2 holds, due to the choice of B_t^* and the concavity of J_{t+1}, we have $G_t''^B(Y_t) \leq 0$ for all Y_t, and therefore $G_t^B(Y_t)$ is concave in Y_t. ■

Proof of Property 3.5

Proof. Let us prove the concavity of $G_t(Y_t)$.

In each period, we must be in one of the following cases; which are independent of the Y_t values:

- If $p_t^2 + r_t^2 - b_t^2 \leq J'_{t+1}(0) \leq p_t^1 + r_t^1 + h_t$, we have $R_t^* = B_t^* = 0$, therefore $R_t^c = B_t^c = 0$, so we have $G_t(Y_t) = G_t^R(Y_t) = G_t^B(Y_t)$, which can be seen from the formulations directly.

- If $J'_{t+1}(0) > p_t^1 + r_t^1 + h_t$, we have $R_t^* \geq 0$ and $B_t^* = 0$, therefore $R_t^c \geq 0$ and $B_t^c = 0$, so we have $G_t(Y_t) = G_t^R(Y_t) \geq G_t^B(Y_t)$.

- If $J'_{t+1}(0) < p_t^2 + r_t^2 - b_t^2$, we have $B_t^* \geq 0$ and $R_t^* = 0$, therefore $B_t^c \geq 0$ and $R_t^c = 0$, so we have $G_t(Y_t) = G_t^B(Y_t) \geq G_t^R(Y_t)$.

We see that in each period, $G_t(Y_t)$ reduces to some function that is proved to be concave. Therefore $G_t(Y_t)$ is concave. ■

Proof of Property 3.6

Proof. The proof is the same as Part 2 in Lemma 3.1 ■

3.9 Appendix C

Experimental Data

Periods	**c**	p^1	p^2	$E(D^1 + D^2)$
1	70	110	90	100
2	90	130	110	70
3	70	110	90	100
4	50	90	70	130
5	70	110	90	100
6	90	130	110	70
7	70	110	90	100
8	50	90	70	130
9	70	110	90	100
10	90	130	110	70
11	70	110	90	100
12	50	90	70	130
Avg	**70**	**110**	**90**	**100**

Table 3.1: Experimental Data for Two-Class-Customer Problem

Chapter 4

Manufacturing Systems with Make-to-Stock and Make-to-Order Products

4.1 Introduction

In recent years, many retail and manufacturing companies have started exploring innovative revenue management techniques in an effort to improve their operations and ultimately the bottom line. Manufacturing systems that can produce different types of products have raised more research interest in recent years as many firms begin to practice market segmentation by providing multiple types of products to customers, and then differentiating customers according to their choices.

Customized production is a strong trend in the manufacturing industry. For example, in the computer industry, many companies allow customers to decide the configuration their products, and companies use the make-to-order (MTO) mode to manage the production. Produced to order production not only gives more satisfaction to customers, but it also helps manufacturers eliminate finished goods inventory. However, the MTO environment suggests important challenges associated with matching fixed production capacity with highly variable demand. Specifically, an MTO system implies periods where the facility is idle and other times in which a large number of orders are awaiting production.

For instance, no company underscores the impact of customized production and customized pricing strategies more than Dell Computers. Dell provides very high flexibility in configurations for customers who are willing to pay high prices. At the same time, Dell also frequently provides promotions for some low-end products to attract more customers, and for these products, customers usually have very little flexibility on configurations. To satisfy these demands, Dell also produces some standard products to stock, the make-to-stock (MTS) environment. This gives raise to a combination of a make-to-stock/make-to-order environment that allows Dell to better manage their production capacity and increase expected profit.

The application of make-to-stock/make-to-order manufacturing systems is also important among part suppliers who face demands from both original equipment manufacturers (OEMs) and the so-called aftermarket. For example, in the automobile industry, a part supplier sells its products to automotive assembly plants for installation into new vehicles, as well as repair shops for replacement in old vehicles (Carr and Duenyas [5]). OEM and aftermarket demands are both important to the part supplier. OEM demands guarantee high utilization of the production capacity, while aftermarket demands bring high profit margins to the supplier. OEM sales are based on long-term contracts, and they are produced under the make-to-stock mode. In contrast, aftermarket items are produced under the make-to-order mode due to their large variety. Note that, as opposed to Dell example, in this case the high priority products are produced to order.

All these developments call for models that integrate production, sequencing and admission decisions for a hybrid production system with both make-to-stock products and make-to-order products. Unfortunately, the academic literature for hybrid manufacturing systems is quite limited.

In this chapter, we study policies to coordinate production, sequencing and admission controls for two types of manufacturing systems with both make-to-stock products and make-to-order products. Our model is different from Carr and Duenyas

[5] in the following sense: (*i*) We study two types of manufacturing systems: in the first case, MTS product has higher priority than MTO product, and in the second case, MTS product has lower priority; While Carr and Duenyas [5] only studied the first case. (*ii*) In our model, unsatisfied high-priority demands are fully backlogged; while in Carr and Duenyas [5], when the high-priority product is out of stock, the demand is lost. Our backlogging assumption is reasonable since it is often difficult to find a product matching all the required features from other suppliers immediately. (*iii*) Our optimal policies are characterized by threshold levels that have a simple linear structure, while in Carr and Duenyas [5], the threshold levels do not have a simple structure.

In the first model of this chapter, we study policies to coordinate production, sequencing, and admission controls for a manufacturing system with both high-priority MTS (e.g., OEM), and low-priority MTO (e.g., aftermarket) demands. High priority customers order identical items from the manufacturer, and these orders cannot be rejected. When the product is out of stock, the demand is fully backlogged and results in penalties for delayed fulfillments. Demands for low-priority, customized products can be rejected if the manufacturer does not have enough production capacity, or otherwise be backordered with a backlogging cost much smaller than that for high-priority orders. This model is similar to the model in Carr and Duenyas [5], except that we assume unsatisfied OEM demands are fully backordered rather than lost. As we will show, this difference has an important impact on the structure of the optimal policy. Indeed, unlike in [5] where the optimal policy is characterized by complex switching curves, in our case the switching curves are linear. In addition, we extend the optimal policy to a problem in which the manufacturer can cancel low-priority orders, as well as to a model with multiple types of MTO products.

In the second model, we examine the the production and admission controls in a manufacturing system with both high-priority MTO (customized) and low-priority MTS (pre-configured) products. MTO products provide high configuration flexibilities to customers, and thus are sold at higher prices, and MTS products are sold at lower prices for promotion. Demands for high-priority MTO products will all be accepted, and the penalties for delayed fulfillments are also higher; while demands for low-priority MTS products will be satisfied if the inventory is available, backordered with a backlogging cost much smaller than high-priority orders, or otherwise rejected if the manufacturer does not have enough production capacity. We characterize the optimal policy with linear threshold levels. We also extend our results to systems with cancelable MTS backorders.

The chapter is organized as follows. In Section 4.2, we study the model with both OEM and aftermarket demands, and we extend it with cancelable aftermarket orders by the manufacturer, as well as with multiple types of aftermarket products. In Section 4.3, we investigate the model with both customized and pre-configured products, and extend the results to systems with cancelable pre-configured backorders. In Section 4.4, we use computational analysis to obtain insights into the benefits of the new policies, and the impact of production capacity, demand structure and cost structure on system performance. Finally we summarize in Section 4.5.

4.2 Model with OEM and Aftermarket Products

4.2.1 Problem Formulation

We consider a manufacturing system with two types of products in an infinite horizon. Type 1 is for high-priority orders and has standard configuration, so it is produced to stock. Demands for type 1 are satisfied if inventory is available, or are fully backlogged with high backlogging penalty b_1 per item per unit time if type 1 is not available in inventory. Type 2 has varying features and is produced to order. Demands for type 2 can be either accepted (backlogged) or rejected. The unit backlogging cost of type 2 is less than that of type 1, $b_2 < b_1$. If a demand for type 2 is rejected, a rejection penalty, r_2, associated with lost sales and loss of good will, is incurred.

We assume that customers pay for products when they place orders rather than when products are received, and we also assume that the production costs are incurred when customers pay for the products, so we can use the selling profit (selling price minus production cost) to substitute the selling price for each class, and we let p_i be the selling profit for Class i, $i = 1, 2$. Finally, we assume linear inventory holding cost, h, per unit time for type 1.

We assume that customers for each type of product arrive according to a Poisson Process, and we let λ_i be the demand arrival rate of Class i, for $i = 1, 2$. Also, we assume that the production time of each type of product follows the same exponential distribution, with production rate μ. We further assume that preemptions are allowed, and no set-up time is needed when the manufacturing system switches from one type of job to the other. This assumption is reasonable for some assembly production systems in which setup times are negligible compared to production times.

We note here that the exponential assumption regarding the production times is what allows us to formulate the problem and characterize the structure of the optimal policy. However, one can simply extend our analysis to systems with phase-type distribution. This makes the analysis more complex while not providing additional insights. Furthermore, after our model reveals the structure of the optimal production and admission control policies, it becomes clear that our insights are not influenced by the assumption on production times.

The system state can be described by a vector of two variables, $\mathbf{y}(t) = (y_1(t), y_2(t))$, where $y_i(t)$ is the net inventory level of product i with $y_1(t) \in Z$ and $y_2(t) \in Z^-$, where Z is the set of all positive and negative integer numbers, while set Z^- only includes non-positive integer numbers. We use $y_1^+(t) = \max\{0, y_1(t)\}$ to show the amount of inventory of type 1 at time t, and $y_1^-(t) = \max\{0, -y_1(t)\}$ to show the number of backorders at time t. Similarly, $-y_2(t)$ is the number of backorders of product 2 at time t. The system state space is $\Omega = Z \times Z^-$.

We use an approach parallel to Ha [17] to analyze the problem. In state $\mathbf{y} = (y_1, y_2)$, the system incurs a cost at rate

$$c(\mathbf{y}) = -hy_1^+ - b_1 y_1^- + b_2 y_2.$$

Let α be the time discount rate, let $N_i^a(t)$ be the number of accepted orders over interval $[0, t]$ for product i, $i = 1, 2$, and let $N_2^r(t)$ be the number of rejected orders for type 2 over the same period of time. We seek an optimal control policy π so as to maximize either the discounted system profit over an infinite horizon,

$$\max_{\pi} \quad J^{\pi}(\mathbf{y}(0)) = \mathrm{E}^{\pi}_{\mathbf{y}(0)}\left[\sum_{i=1}^{2}\int_0^{\infty} e^{-\alpha t}p_i dN_i^a(t) - \int_0^{\infty} e^{-\alpha t}r_2 dN_2^r(t) + \int_0^{\infty} e^{-\alpha t}c(\mathbf{y}(t))dt\right], \tag{4.1}$$

or the average profit over an infinite horizon,

$$\max_{\pi} \quad J_a^{\pi} = \lim_{T\to\infty}\frac{1}{T}\mathrm{E}^{\pi}\left[\sum_{i=1}^{2} p_i N_i^a(T) - r_2 N_2^r(T) + \int_0^T c(\mathbf{y}(t))dt\right]. \tag{4.2}$$

In (4.1), $J_{\pi}(\mathbf{y}(0))$ is the expected profit function under policy π starting from initial state $\mathbf{y}(0) = (y_1(0), y_2(0))$. In the rest of this chapter we will mainly focus on the discount-profit problem. However, the theoretical results in the discounted-profit model also apply to the average-profit problem.

The optimal function $J^*(y_1, y_2)$ satisfies the following optimality equation [3]:

$$J(y_1,y_2) = \frac{1}{\alpha+\mu+\lambda_1+\lambda_2}\Big\{c(y_1,y_2)+\mu H_0 J(y_1,y_2)+\lambda_1 H_1 J(y_1,y_2)+\lambda_2 H_2 J(y_1,y_2)\Big\} \tag{4.3}$$

where H_0, H_1, and H_2 are functions defined by,

$$\begin{aligned} H_0 J(y_1,y_2) &= \max\Big\{J(y_1,y_2), J(y_1+1,y_2), J(y_1,y_2+1|y_2<0)\Big\} \\ H_1 J(y_1,y_2) &= J(y_1-1,y_2)+p_1 \\ H_2 J(y_1,y_2) &= \max\Big\{J(y_1,y_2-1)+p_2, J(y_1,y_2)-r_2\Big\}. \end{aligned}$$

H_0 corresponds to the production decision: the manufacturer can choose to either produce or stop production. As a constraint, MTO products can only be produced when there are backorders. H_1 indicates that the demands for type 1 will always be accepted. H_2 is associated with the admission control for an arriving demand for type 2. The manufacturer can either accept (backlog) or reject the demand.

Since it is always possible to redefine the time scale, without loss of generality, we can assume $\alpha+\mu+\lambda_1+\lambda_2 = 1$. Then, the optimality equation can be simplified as:

$$J(y_1,y_2) = c(y_1,y_2)+\mu H_0 J(y_1,y_2)+\lambda_1 H_1 J(y_1,y_2)+\lambda_2 H_2 J(y_1,y_2) := HJ(y_1,y_2) \tag{4.4}$$

The optimality equation under the average-profit criterion is:

$$J(y_1, y_2) + g = c(y_1, y_2) + \mu H_0 J(y_1, y_2) + \lambda_1 H_1 J(y_1, y_2) + \lambda_2 H_2 J(y_1, y_2) \qquad (4.5)$$

where g is the optimal average profit per unit time.

4.2.2 The Optimal Policy

We investigate the structure of the optimal policy following the approach of Ha [18] and De Vericourt, Karaesmen and Dallery [8]. We first define a set of optimality conditions and decision rules, and then show that the optimal expected profit function, $J(\mathbf{y})$, satisfies the conditions.

For any function f defined on Ω, let $\Delta_1 f(y_1, y_2) = f(y_1+1, y_2) - f(y_1, y_2)$, let $\Delta_2 f(y_1, y_2) = f(y_1, y_2+1) - f(y_1, y_2)$, and let $\Delta_{12} f(\mathbf{y}) = f(y_1+1, y_2) - f(y_1, y_2+1)$. We define the set of functions as $\mathcal{C}$, such that if $f(y_1, y_2) \in \mathcal{C}$, then,

- **C.1:** For $(y_1, y_2) \in \Omega$,

  Condition C.1.1: $\Delta_i f(y_1, y_2) \geq 0$, if $y_i < 0$, $i = 1, 2$;

  Condition C.1.2: $\Delta_{12} f(y_1, y_2) \geq 0$, if $y_1 < 0$.

- **C.2:** For $(y_1, y_2) \in \Omega$,

  Condition C.2.1: $\Delta_1 f(y_1, y_2)$ and $\Delta_2 f(y_1, y_2)$ are non-increasing in y_1 and y_2;

  Condition C.2.2: $\Delta_1 f(y_1, y_2) \geq 0$ for $y_1 < S$, where $S = \min\{z | \Delta_1 f(z, 0) < 0\}$.

- **C.3:** For $(y_1, y_2) \in \Omega$ and $y_2 < 0$,

  Condition C.3.1: $\Delta_{12} f(y_1, y_2)$ is non-increasing in y_1 and independent of y_2;

  Condition C.3.2: $\Delta_{12} f(y_1, y_2) \geq 0$ for $y_1 < R$, where $R = \min\{z | \Delta_{12} f(z, -1) < 0\}$.

- **C.4:** For $(y_1, y_2), (y_1', y_2') \in \Omega$ and $y_2 < 0$,

  Condition C.4.1: $\Delta_2 f(y_1, y_2) = \Delta_2 f(y_1', y_2')$, if $y_1 + y_2 = y_1' + y_2'$;

  Condition C.4.2: $\Delta_2 f(y_1, y_2 - 1) \leq p_2 + r_2$ for $y_1 + y_2 > B$, where $B = \max\{z | \Delta_2 f(0, z-1) > p_2 + r_2\}$.

To have some intuition on the above conditions, we apply the condition set $\mathcal{C}$ to the expected profit function $J(y_1, y_2)$. Condition **C.1.1** implies that, if there are backorders for a product, it is better to produce the product rather than idle the machine. Condition **C.1.2** indicates that, if there are backorders for both products, type 1 has higher priority than type 2. Condition **C.2.1** implies that the marginal benefits of increasing y_1 and y_2 are non-increasing in both y_1 and y_2. Condition **C.2.2** indicates that for threshold level S, if $y_1 < S$, then producing type 1 is better than idling the machine. Condition **C.3.1** and **C.3.2** imply that type 1 has higher priority if the inventory of product 1 is less than a threshold level R (i.e., $y_1 < R$), and type 2 has higher priority otherwise. The sign of $p_2 + r_2 - \Delta_2 J(y_1, y_2 - 1)$ determines whether to reject an order for type 2. Condition **C.4.1** suggests that admission decisions depend on total inventory level, but not on the inventory level of each product, and Condition **C.4.2** indicates the acceptance of a new order for type 2 rather than rejection, as long as the the total inventory of products 1 and 2 are larger than a threshold level B (i.e., $y_1 + y_2 > B$).

We show that under the optimal conditions, the threshold level S cannot be less than threshold level R.

Proposition 4.1 *In the optimal solution $R \leq S$.*

In the following, we show that the optimal profit function $J(y_1, y_2)$ satisfies all the conditions in set $\mathcal{C}$ (i.e., $J(y_1, y_2) \in \mathcal{C}$). Lemma 4.1 indicates that the structure of the function f in $\mathcal{C}$ is preserved under functions H.

Lemma 4.1 *If $f(y_1, y_2) \in \mathcal{C}$, then $Hf(y_1, y_2) \in \mathcal{C}$.*

In the next theorem we use the lemma to show that a modified basestock policy is optimal.

Theorem 4.1 *The optimal policy is characterized by three parameters: the basestock level, S, the rationing level, R, and the admission level, B, such that,*

- *Production control: when there are no backorders in the system, it is optimal to produce type 1 if $y_1 < S$, and to stop production if $y_1 \geq S$.*
- *Rationing control: when there are backorders in the system, it is optimal to produce type 1 if $y_1 < R$, and to produce type 2 if $y_1 \geq R$ and $y_2 < 0$.*
- *Admission control: it is optimal to accept an arriving demand for type 2 if $y_1 + y_2 > B$, and to reject otherwise.*

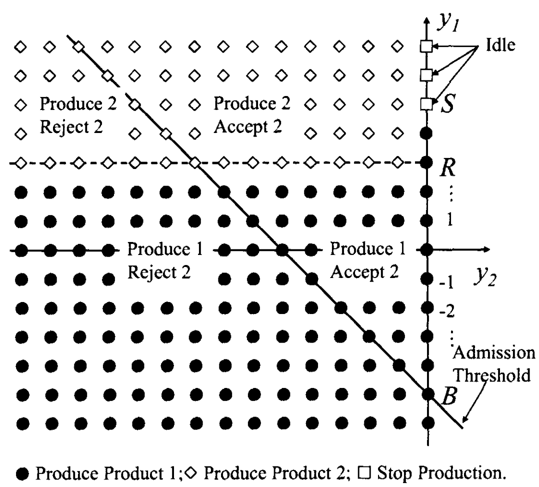


Figure 4-1: The (S, R, B) Policy

The optimal policy can be described in Figure 4-1. Under the optimal policy, the manufacturer will stop production if the inventory level of type 1 is greater or equal to S and there is no backorder for type 2. When the inventory level of type 1 is higher than R and there are backorders of type 2 in the system, the manufacturer will give higher priority to type 2 backorders. If the inventory level of type 1 is less than R or there are no backorders for type 2, the manufacturer will produce type 1 to increase the inventory level. The manufacturer will accept an arriving demand for type 2 if the total net inventory $y_1 + y_2$ is higher than B or reject it otherwise. We refer to this policy as the (S, R, B) policy.

4.2.3 Cancelable Aftermarket Backorders

In this subsection, we extend our results to the problem where backorders for MTO products can be canceled by the manufacturer. We assume that the manufacturer can cancel a backorder for type 2 with a cancelation penalty l_2, $l_2 \geq r_2$. In this case, the manufacturer needs to decide whether to cancel a low-priority backorder as a high-priority order arrives. We first change function H_1 as

$$H_1 J(y_1, y_2) = \max\Big\{J(y_1 - 1, y_2) + p_1, J(y_1 - 1, y_2 + 1) + p_1 - p_2 - l_2\Big\}.$$

Then we can add the following condition to $\mathcal{C}$:

- **C.4:** For $(y_1, y_2) \in \Omega$, $y_2 < 0$,

  Condition C.4.3: $\Delta_2 f(y_1, y_2 - 1) \leq p_2 + l_2$ for $y_1 + y_2 > L$, where

$$L = \max\{z | \Delta_2 f(0, z - 1) > p_2 + l_2\}$$

.

We show that under the optimal conditions, the threshold level B cannot be greater than threshold level L.

Proposition 4.2 *In the optimal solution $L \leq B$.*

With a few modifications in the proofs, we can show that Lemma 1 holds with the modified functions and the updated condition set. In addition, we have the following theorem:

Theorem 4.2 *The optimal policy is characterized by four parameters: the basestock level, S, the rationing level, R, the admission level, B, and the cancelation level, L, such that,*

- *Production control: when there are no backorders in the system, it is optimal to produce type 1 if* $y_1 < S$*, and to stop production if* $y_1 \geq S$*.*

- *Rationing control: when there are backorders in the system, it is optimal to produce type 1 if* $y_1 < R$*, and to produce type 2 if* $y_1 \geq R$ *and* $y_2 < 0$*.*

- *Admission control: it is optimal to accept an arriving demand for type 2 if* $y_1 + y_2 > B$*, and to reject otherwise.*

- *Cancelation control: it is optimal to cancel a backorder for type 2 upon arrival of an order for type 1, if* $y_1 + y_2 \leq L$*.*

The optimal policy for a two-product problem with cancelation is described in Figure 4-2. Under the optimal policy, if the total inventory $y_1 + y_2$ is lower or equal to L, the manufacturer will cancel a backorder for type 2 when an order for type 1 arrives. We refer to this policy as the (S, R, B, L) policy.

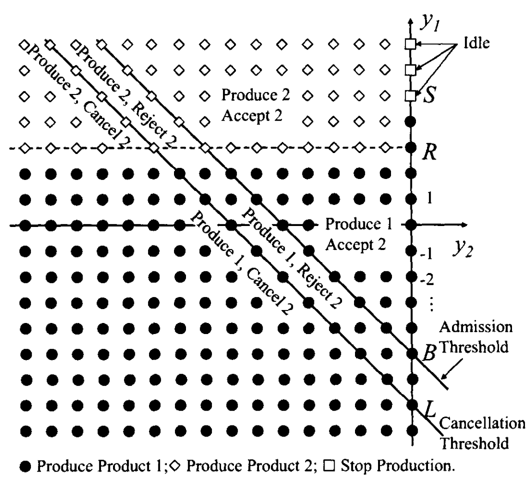


Figure 4-2: The (S, R, B, L) Policy

4.2.4 Multiple Make-to-Order Products

Using the same approach as in Carr and Duenyas [5], we can extend the above model to a problem with a single MTS product and multiple MTO products. For this purpose, we use the same assumption as in Carr and Duenyas [5] that all MTO products have the same production rate and the same backlogging penalty. Without loss of generality, the aftermarket products can be indexed from 2 to n, such that $p_2 + r_2 \leq \ldots \leq p_n + r_n$.

Since all MTO demands have the same backlogging penalty, it is unnecessary to differentiate them once they are accepted. All MTO demands have the same priority in the waiting queue, and the first-come-first-serve rule will be applied. So the system state can still be described with (y_1, y_2), where y_1 is the net inventory level of the OEM product, and $-y_2$ is the total amount of backorders of all MTO products (product 2 to product n).

To adjust the model to this problem, we change the function H_i, $i = 2, \ldots, n$, as follows:

$$H_i J(y_1, y_2) = \max\Big\{J(y_1, y_2 - 1) + p_i, J(y_1, y_2) - r_i\Big\}.$$

Then we can change condition **C.4.2** to,

- **C.4:** For $(y_1, y_2) \in \Omega$, $y_2 < 0$,

  Condition C.4.2: $\Delta_2 f(y_1, y_2 - 1) \leq p_i + r_i$ for $y_1 + y_2 > B_i$, where

$$B_i = \max\{z | \Delta_2 f(0, z - 1) > p_i + r_i\}$$

.

Proposition 4.3 *In the optimal solution $B_i \geq B_j$, for $2 \leq i < j$.*

With some modifications in the proof, we can show that Lemma 1 holds for the modified function, H_i, $i = 2, \ldots, n$, and the updated condition set. The optimal policy of the problem with multiple aftermarket products can be characterized by the basestock level, S, the rationing level, R, and a sequence of admission levels, $B_2, \ldots, B_n$, where B_i, $i = 2, \ldots, n$, is the admission level of product i.

The optimal policy for a three-product problem with cancelation is described in Figure 4-3. If the total inventory, $y_1 + y_2$, is between B_2 and B_3, the manufacturer will reject demands for type 2, and accept demands for type 3; If the total inventory is below B_3, the manufacturer will reject demands both for type 2 and 3.

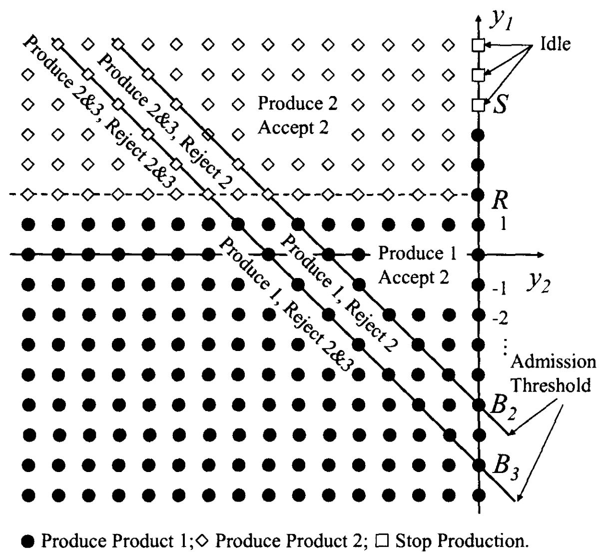


Figure 4-3: The (S, R, B) Policy with Two Types of Make-to-Order Products

4.3 Model with Customized and Pre-configured Products

4.3.1 Problem Formulation

In this section, we consider another type of MTS/MTO manufacturing system, which pervades in the companies facing consumers directly. For example, Dell provides very high flexibility for customers to customize their products, and at the same time, it also frequently provide promotions for some low-end, pre-configured products. The promotional products not only enlarge the company's market share, but also allows the manufacturer to better manage the production capacity.

In this model, type 1 is for promotion, and it is produced to stock; type 2 is the regular product for customization, and hence it is produced to order. Demands for type 1 is satisfied if inventory is available, and is backlogged or rejected otherwise depending on the manufacturer decision. Demands for type 2 provide higher profit margin to the manufacturer, and they are always fully accepted. Type 1 has lower unit backlogging penalty than type 2, $b_1 \leq b_2$. We continue with the notations and the assumptions in Section 2.

With the above assumptions, we observe that the manufacture does not control the production and sequencing of type 2 demands. Orders for type 2 always have higher priorities than for type 1, and within each class, the FCFS policy is respected. The expected profit generated by type 2 product is determined by the demand process itself, and the manufacturer only maximizes the "extra" profit generated by the promotional product. The production and admission control of type 1 is independent of the profit and backlogging penalty of type 2, as long as $b_1 \leq b_2$. Therefore, the two-product problem is reduced to a single-product problem under the impacts of the higher-priority product, and we only need to incorporate the unit profit and the backlogging penalty of type 1 into the model. For the same reason, this problem can be easily extended with $n-1$ types of MTO products whose backlogging penalties, $b_2, \ldots, b_n$ are all higher than b_1. Orders for different types of MTO products are sequenced according to their backlogging penalties, and their impacts on the MTS product are independent of their types.

In the following, we use an approach parallel to Section 4.2 to analyze a system with one MTS and one MTO product. In state $\mathbf{y} = (y_1, y_2)$, the system incurs an "extra" cost for type 1 at rate

$$c_1(y_1) = -hy_1^+ - b_1 y_1^-.$$

We seek an optimal control policy π so as to maximize either the discounted "extra" profit generated by type 1 over an infinite horizon,

$$\max_{\pi} \quad J_1^{\pi}(\mathbf{y}(0)) = E_{\mathbf{y}(0)}^{\pi}\left[\int_0^{\infty} e^{-\alpha t} p_1 dN_1^a(t) - \int_0^{\infty} e^{-\alpha t} r_1 dN_1^r(t) + \int_0^{\infty} e^{-\alpha t} c_1(\mathbf{y}(t)) dt\right], \tag{4.6}$$

or the average "extra" profit generated by type 1 over an infinite horizon,

$$\max_{\pi} \quad J_{1a}^{\pi} = \lim_{T \to \infty} E^{\pi}\left[p_1 N_1^a(T) - r_1 dN_1^r(T) + \int_0^T c_1(\mathbf{y}(t)) dt\right], \tag{4.7}$$

In (4.6), $J_1^{\pi}(\mathbf{y}(0))$ is the expected profit generated by type 1 under policy π starting from initial state $\mathbf{y}(0) = (y_1(0), y_2(0))$. In the rest, we will mainly focus on the discounted-profit problem. The theoretical results in the discounted-profit problem are also applicable in the average-profit model.

It can be shown that the optimal function $J_1^*(y_1, y_2)$ satisfies the following optimality equation [3]:

$$J_1(y_1,y_2) = \frac{1}{\alpha+\mu+\lambda_1+\lambda_2}\Big\{c_1(y_1)+\mu H_0 J_1(y_1,y_2)+\lambda_1 H_1 J_1(y_1,y_2)+\lambda_2 H_2 J_1(y_1,y_2)\Big\} \tag{4.8}$$

where H_i, $i = 0, 1, 2$, are functions defined in Ω, such that,

$$\begin{aligned} H_0 J_1(y_1,y_2) &= \max\Big\{J_1(y_1,y_2), J_1(y_1+1,y_2), J_1(y_1,y_2+1|y_2<0)\Big\} \\ H_1 J_1(y_1,y_2) &= \max\Big\{J_1(y_1,y_2)-r_1, J_1(y_1,y_2-1)+p_1\Big\} \\ H_2 J_1(y_1,y_2) &= J_1(y_1,y_2-1) \end{aligned}$$

H_0 corresponds to the result of the production decision. The manufacturer can choose to either produce or stop production. Particularly, make-to-order products can only be produced when there are backorders. H_1 is associated with the admission control for a new arrival of type 1 demand. The manufacturer will satisfy demand if inventory is available, or choose to backlog or to reject otherwise. Function H_2 indicates that an arriving demand for type 2 will always be accepted.

Again, we can assume $\alpha + \mu + \lambda_1 + \lambda_2 = 1$ by redefining the time scale, and the optimality equation can be simplified as:

$$J_1(y_1,y_2) = c_1(y_1)+\mu H_0 J_1(y_1,y_2)+\lambda_1 H_1 J_1(y_1,y_2)+\lambda_2 H_2 J_1(y_1,y_2) := HJ_1(y_1,y_2) \tag{4.9}$$

The optimality equation under the average-profit criterion is:

$$J_1(y_1,y_2)+g_1 = c_1(y_1)+\mu H_0 J_1(y_1,y_2)+\lambda_1 H_1 J_1(y_1,y_2)+\lambda_2 H_2 J_1(y_1,y_2) \tag{4.10}$$

where g_1 is the optimal average profit generated by type 1 per unit time.

4.3.2 The Optimal Policy

We use the same approach in Section 4.2.2 to investigate the structure of the optimal policy for this model. Let us define a set $\mathcal{D}$ of functions such that if $f(y_1, y_2) \in \mathcal{D}$, then,

- **D.1:** For $(y_1, y_2) \in \Omega$,

  Condition D.1.1: $\Delta_i f(y_1, y_2) \geq 0$, if $y_i < 0$, $i = 1, 2$.

  Condition D.1.2: $\Delta_{12} f(y_1, y_2) \leq 0$.

- **D.2:** For $(y_1, y_2) \in \Omega$,

  Condition D.2.1: $\Delta_1 f(y_1, y_2)$ is non-increasing in y_1 and y_2.

  Condition D.2.2: $\Delta_1 f(y_1, y_2) \geq 0$ for $y_1 < S$, where $S = \min\{z | \Delta_1 f(z, 0) < 0\}$.

- **D.3:** For $(y_1, y_2), (y_1', y_2') \in \Omega$, $y_1 < 0$, $y_1' < 0$,

  Condition D.3.1: $\Delta_1 f(y_1, y_2) = \Delta_1 f(y_1', y_2')$, if $y_1 + y_2 = y_1' + y_2'$.

  Condition D.3.2: $\Delta_1 f(y_1 - 1, y_2) \leq p_1 + r_1$ for $y_1 + y_2 > B$, where

$$B = \max\{z | \Delta_1 f(z - 1, 0) > p_1 + r_1\}$$

.

To get some intuition on the above conditions, we apply the condition set $\mathcal{D}$ to the expected profit, $J_1(y_1, y_2)$. Condition **D.1.1** implies that if there are backorders in any class, the production line needs to keep producing. Condition **D.1.2** indicates that a backorder for type 2 always has a higher priority than for type 1. Condition **D.2.1** implies that the marginal benefit of increasing y_1 is non-increasing in both y_1 and y_2. Condition **D.2.2** indicates that if $y_1 < S$, then the production line needs to keep producing type 1 when there are backorders for type 2. Since the sign of $p_1 + r_1 - \Delta_1 J_1(y_1 - 1, y_2)$ determines whether to reject an order for type 1, Condition **D.3.1** suggests that the rejection decisions depend on total inventory level. Condition **D.3.2** indicates to accept an arriving demand for type 1 as long as the level of inventory satisfies $\sum_{i=1}^{n} y_i > B$.

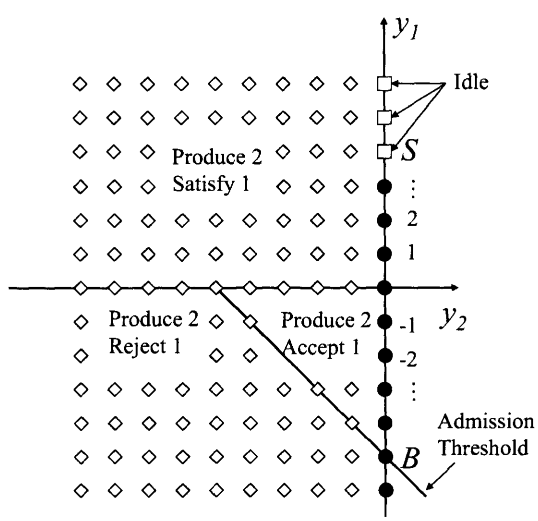
The following lemma shows that the structure of the functions in $\mathcal{D}$ is preserved under functions H.

Lemma 4.2 *If $f(y_1, y_2) \in \mathcal{D}$, then $Hf(y_1, y_2) \in \mathcal{D}$.*

Using the lemma we show, in the next theorem, that a modified basestock policy is optimal.

Theorem 4.3 *The optimal policy is characterized by two parameters: the base-stock level, S, and the admission level, B, such that,*

- *Production control: if there are backorders for type 2, it is optimal to produce type 2; otherwise, it is optimal to produce type 1 if $y_1 < S$, and to stop production if $y_1 \geq S$;*
- *Admission control: it is optimal to satisfy an arriving demand for type 1 if $y_1 > 0$, to reject if $y_1 + y_2 \leq B$, and to backlog otherwise.*



● Produce Product 1;◇ Produce Product 2; □ Stop Production.

Figure 4-4: The (S, B) Policy

The optimal policy is described in Figure 4-4. Under the optimal policy, the manufacturer will stop production if the inventory level of type 1 reaches S, and keep producing otherwise. When there is no inventory available, the manufacturer will backlog an arriving demand for type 1 if the net inventory, $y_1 + y_2$, is higher than B, or reject otherwise. In this policy, type 2 orders always have higher priority and cannot be rejected. We refer to this policy as the (S, B) policy.

4.3.3 Cancelable Promotional Backorders

In this section, we assume the manufacturer can cancel a backorder for type 1 with a cancelation penalty l_1, $l_1 \geq r_1$. In this case, the manufacturer needs to decide whether to cancel a backorder for type 1 when an order for product 2 arrives. To adjust the model to this situation, we first change function H_2 as follows

$$H_2 J_1(y_1, y_2) = \max\Big\{J_1(y_1, y_2 - 1), J_1(y_1 + 1, y_2 - 1) - p_1 - l_1\Big\}.$$

Then we can add the following condition to $\mathcal{D}$:

- **D.3:** For $(y_1, y_2) \in \Omega$, $y_1 < 0$,

  Condition D.3.3: $\Delta_1 f(y_1 - 1, y_2) \leq p_1 + l_1$ for $y_1 + y_2 > L$, where

$$L = \max\{z | \Delta_1 f(z - 1, 0) > p_1 + l_1\}.$$

We first show that under the optimal conditions, the threshold level B cannot be greater than L.

Proposition 4.4 *In the optimal solution $L \leq B$.*

With a few modifications in the proof, we can show that Lemma 4.2 holds with the modified functions and the updated conditions. In addition, we have the following theorem:

Theorem 4.4 *The optimal policy for the problem with cancellable make-to-stock backorders is characterized by three parameters: the base-stock level, S, the admission level, B, and the hold-up-to level, L, such that,*

- *Production control: if there are backorders for type 2, it is optimal to produce type 2; otherwise, it is optimal to produce type 1 if $y_1 < S$, and to stop production if $y_1 \geq S$;*
- *Admission control: it is optimal to satisfy an arriving demand for type 1 if $y_1 > 0$, to reject if $y_1 + y_2 \leq B$, and to backlog otherwise;*
- *Cancelation control: it is optimal to cancel a backorder for type 1 after accepting an order for a higher-priority product, if $y_1 + y_2 \leq L$.*

The optimal policy is described in Figure 4-5. Under the optimal policy, if the net inventory level, $y_1 + y_2$, is lower or equal to L, the manufacturer will cancel a backorder for type 1 when an order for type 2 arrives. We refer to this policy as the (S, B, L) policy.

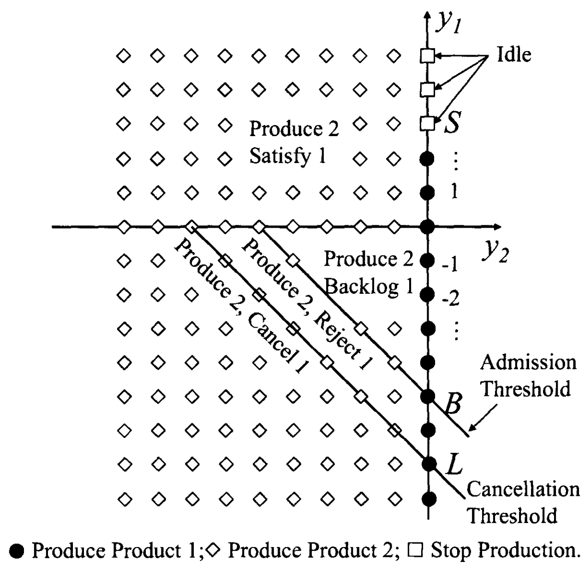


Figure 4-5: The (S, B, L) Policy

4.4 Computational Analysis

In this section we report the results of the computational study that we performed based on the model with OEM and aftermarket products. The computational analysis based on the model with customized and pre-configured products presents similar properties, and therefore is omitted.

The goal of our numerical analysis is to investigate the following questions:

1. What is the profit improvement of the optimal policy relative to commonly used policies? Under what circumstances the improvement is or is not very significant?

2. Canceling orders prevents large backlogging costs for the manufacturer. However, it also decreases customer satisfaction. How much the profit decreases if the manufacturer does not utilize the flexibility of canceling the customer orders?

3. In some systems the production capacity is separated between the two products in a dedicated mode. Pooling the capacity dedicated to the two products and creating a flexible capacity may be very costly. Thus, it is of interest to study in what circumstances pooling the capacity of the two products provides a significant increase in profit.

4. When the objective is to maximize profit, it is sometimes inevitable to sacrifice the service level of some customers. Therefore, an important question is: What is the impact of the (S, R, B) policy on the service level of both the aftermarket and the OEM customers? Does system with pooled capacity also provide a better service level than systems in which the production capacity is allocated to two products? How and how much will the service level of each product be affected by the pooling effect?

Our numerical study consists of 160 cases generated by varying the following parameters:

- $\rho = (\lambda_1 + \lambda_2)/\mu$: This parameter is an indication of the manufacturer's relative capacity compared to the market size. We refer to ρ as *potential machine utilization* in the rest of the chapter. Please notice that ρ is not the real machine utilization since aftermarket demands may be rejected. We considered five values for ρ, specifically, $\rho \in \{0.6, 0.8, 1.0, 1.2, 1.4\}$.

- $\lambda_1/(\lambda_1+\lambda_2)$: This parameter represents the size of the OEM demand compared to the aftermarket demand. In our numerical study we considered four cases, namely $\lambda_1/(\lambda_1 + \lambda_2) \in \{0, 0.3, 0.7, 1.0\}$.

- p_2/p_1: This ratio represents the price difference between the two products. Since the price of aftermarket product is higher than the price offered to the OEM, we consider four values for p_2/p_1 that are all greater than one, namely $p_2/p_1 \in \{1.0, 1.3, 1.6, 2.0\}$. We did not consider ratios grater than 2, since it is very uncommon in practice to have cases in which an item is sold twice of its price for OEM.

- The penalty ratio, b_2/b_1, represents the backlogging penalty difference between the two products. In the numerical study we assume the backlogging penalty of each product is linearly related to its price, and we consider the following two scenarios: ($b_1 = 20\%p_1, b_2 = 5\%p_2$), and ($b_1 = 40\%p_1, b_2 = 5\%p_2$). Note that, although $p_1 < p_2$, in both scenarios backlogging of type 1 customers is costlier than that of type 2 customers.

In order to better present the effects of the parameters on the system performance (and omit the effects of discount factor), our numerical study uses total expected profit per unit time as the performance measure.

4.4.1 Simple Base-Stock Policy Versus Optimal Policy

In this section we compare the performance of the optimal policy with a policy that is commonly used in practice (i.e., the simple base-stock policy). The objective is to investigate how much the expected profit increases if systems switch from using the simple base-stock policy to using the optimal policy.

The policy that is often used in practice is characterized by two threshold levels: base-stock level, S', for type 1 products, and admission level, B', for type 2 products. Under this policy, the system follows a base-stock policy for type 1 products, namely the system produces type 1 products as long as the inventory is less than S'. When the inventory of type 1 product reaches S', the system produces type 2 products if there is a backorder of type 2, or idles otherwise. The orders for type 2 product will only be accepted if the number of type 2 backorders is less than B'. We call this policy the (S', B') policy.

In order to obtain the system's profit under the (S', B') policy we revised the MDP model (4.3) as follows: For every state of the system we only allow one action, namely the action that is consistent with the (S', B') policy. This makes our MDP model a successive approximation model that results in the average profit under per unit time if policy (S', B') is implemented.

To obtain the optimal values for S' and B', we searched all possible combinations for values S' and B' and obtained (S'^*, B'^*) that results in the maximum expected profit. Considering J_{SB} as the system's profit under the optimal (S', B') policy, we evaluate the profit improvement under the optimal policy using the following measure:

$$Profit\ Potential_{optimal} = \frac{J_{SRB} - J_{SB}}{J_{SB}} \times 100\%$$

where J_{SRB} is the profit under the optimal (S, R, B) policy.

Based on our numerical study, we found that, on average, the profit improvement obtained by using the optimal policy is 8%. The maximum profit improvement can be up to 40%.

We examined the effects of price ratio and potential machine utilization on the profit improvement. Figure 4-6, which is one set of examples among several that we studied, depicts the typical behavior of the system. Figure 4-6 is for a case in which $\lambda_1/(\lambda_1 + \lambda_2) = 0.7$ and $b_1 = 20\% p_1$. As the figure shows, when the price ratio increases, implementing the optimal policy results in more profit improvement. The reason is that, when the price ratio increases, it becomes more beneficial to produce type 2 products. Therefore, the system tends to give more priority to produce type 2 products. On the other hand, since the backlogging penalty cost for type 1 products is higher than that for type 2 products, the system also has the tendency to give priority to type 1 products. Under these circumstances, where the two products are competing for the capacity, the optimal policy can manage the production and the inventory more effectively than the simple base-stock policy.

The figure also suggests that profit potential increases as ρ increases. This is true since as potential machine utilization increases (i.e., capacity becomes tighter) it is more and more important to allocate capacity effectively between the two classes of products. This is exactly what the optimal policy achieves. However, when $\rho > 1.0$, the expected profits under both the optimal policy and the simple base-stock policy become negative as ρ further increases, and the profit difference between the two policies becomes smaller and finally approaches zero. In this case, since the production capacity is very tight, both policies behave the same and allocate all production capacity to type 1 products, and reject the aftermarket demands.

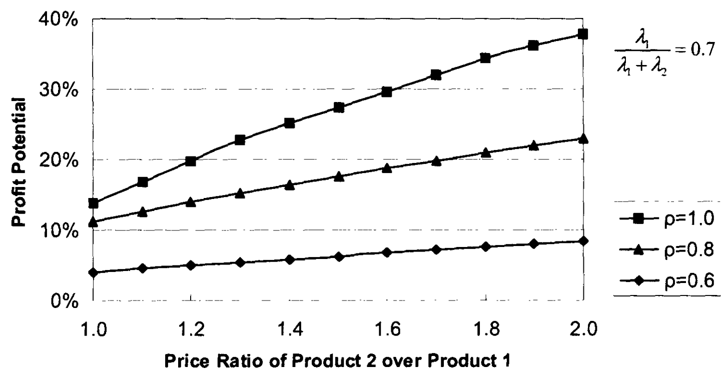


Figure 4-6: Impact of Price Ratio on Profit Potential.

Figure 4-7 shows the typical system behavior of how changes in demand ratio and potential machine utilization affect the profit improvement. Figure 4-7 corresponds to one set of our numerical study in which $p_2/p_1 = 1.6$ and $b_1 = 20\% p_1$. As the figure suggests, when the demand ratio, $\frac{\lambda_1}{\lambda_1+\lambda_2}$, equals 0 or 1, there is only one type of products in the system, and therefore there is no difference between the optimal policy and the simple base-stock policy. However, as the demand ratio increases from 0 to 1, the relative performance of the optimal policy increases and then decreases. The reason is the same as what was described for Figure 4-6. When the demand ratio is not close to 0 or 1, both products are competing for production capacity, and therefore employing the optimal policy that effectively allocates the production capacity between the two products can dramatically improve the profit.

Figure 4-7 also depicts that, as the total demand relative to the production capacity increases, the profit improvement of the optimal policy increases. In other words, manufacturers with tighter capacity benefit more form implementing the optimal policy. The reason is the same as for Figure 4-6.

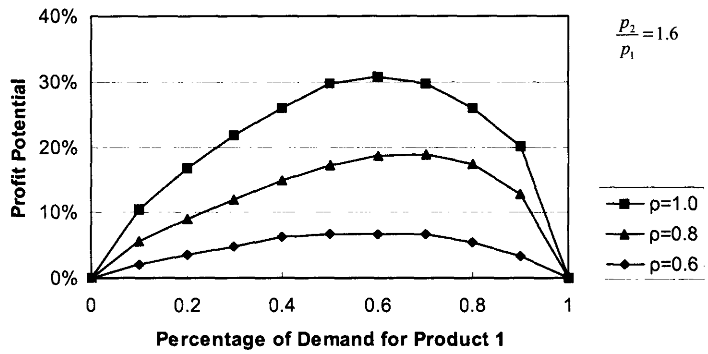


Figure 4-7: Impact of Demand Ratio

4.4.2 Effectiveness of Order Cancelation

We also studied the performance of the cancelation policy analyzed in Section 4.2.3. Interestingly, we found that in all the scenarios in our numerical study, cancelation provides very slight profit increase (i.e., less than 0.8% in maximum), compared with the optimal policy in which order cancelation is not applied. This is quite intuitive. When the manufacturer has the flexibility to accept or reject an order, having additional flexibility of canceling an order later does not have much value. The small amount of profit increase together with the concerns associated with loss-of-reputation explains why cancelation policies are rarely used in supply chain practice.

4.4.3 Effectiveness of Pooling

As we mentioned, pooling the capacity dedicated to the two products and creating a flexible capacity may be very costly. Thus, it is of interest to study in what circumstances pooling the capacity of the two products provides a significant increase in profit.

In this section we therefore study the pooling effect by comparing the performance of a single-machine system with capacity μ with that of a two-machine system with total capacity μ. In the two-machine model, each machine is dedicated to one product. The total production capacity μ is optimally allocated between the two products. Specifically:

- Machine 1 produces item 1 according to a make-to-stock policy with base-stock level S_1. This machine has the capacity to produce μ_1 items of type 1 per unit time.

- Machine 2 produces item 2 according to a make-to-order routine with admission level B_2. Machine 2 has the capacity to produce μ_2 items of type 2 per unit time, where $\mu_1 + \mu_2 = \mu$.

In order to obtain the optimal allocation of capacity μ to machines 1 and 2 (i.e., finding μ_1^* and μ_2^*), we compared the summation of the average profits of both machines for all possible values of (μ_1, μ_2). For each capacity allocation (μ_1, μ_2) we obtained the optimal base-stock level S_1^* that results in the maximum profit for machine 1 and the optimal admission level B_2^* that results in the maximum profit for machine 2. By searching over all capacity allocation (μ_1, μ_2) (and its corresponding optimal threshold level S_1^* and B_2^*), we obtained the optimal capacity allocation (μ_1^*, μ_2^*) that resulted in the maximum total average profit. We call this profit J_{Two}. We then measured the profit potential of polling capacity by:

$$Profit\ Potential_{pooling} = \frac{J_{SRB} - J_{Two}}{J_{Two}} \times 100\%$$

By examining the 160 cases in on our numerical study, we found that, pooling capacity of the two machines and implementing the optimal (S, R, B) policy in the new system can improve profit, on average, by 20%. The maximum improvement can be up to 80%.

We also observed that the potential profit of pooling capacity increases as the price ratio p_2/p_1 or potential machine utilization increases. However, as production capacity gets tighter (e.g., $\rho > 1$), the profit potential beginning to decreases as ρ further increases. Furthermore, similar to the previous case, the potential profit of pooling is low when one demand type is much lower (or higher) than the other (i.e., $\lambda_1/(\lambda_1 + \lambda_2)$ is either very small or very large).

4.4.4 Impact on Service Level

When the objective is to maximize profit, it is sometimes inevitable to sacrifice the service level of some customers. Therefore, it is important to understand what the impact of the (S, R, B) policy is on the service level of both the aftermarket and the OEM customers. Thus, in this section we study the impact of the optimal policy on customers' service levels.

Here we define the service level for aftermarket demands, SL_{AMK}, as the *fraction of aftermarket orders that are accepted*; and the service level for OEM demands, SL_{OEM}, as the *fraction of OEM orders that are immediately satisfied.*

Figure 4-8 illustrates service level for aftermarket demands under the optimal (S, R, B) policy and the simple (S', B') policy. As the figure shows, when the price ratio increases, both the optimal policy and the simple base-stock policy increase the service level for aftermarket demands slightly. This is very intuitive. The manufacturer will accept more aftermarket demands when they contribute more profits.

The figure also suggests more service level improvement as ρ increases. When the production capacity is sufficient, e.g., $\rho = 0.6, 0.8$, both policies maintain high service levels for aftermarket demands, and the service level under the optimal policy is slightly higher than that under the simple base-stock policy. As capacity becomes tight ($\rho = 1.0$), the optimal policy provides higher service level than the simple base-stock policy. The reason is the same as that described in Figure 4-6. However, as ρ further increases, the difference in service levels of the two policies becomes smaller and finally approaches zero. In this case, since the production capacity is very tight, both policies behave the same and allocate all production capacity to type 1 products, and reject the aftermarket demands.

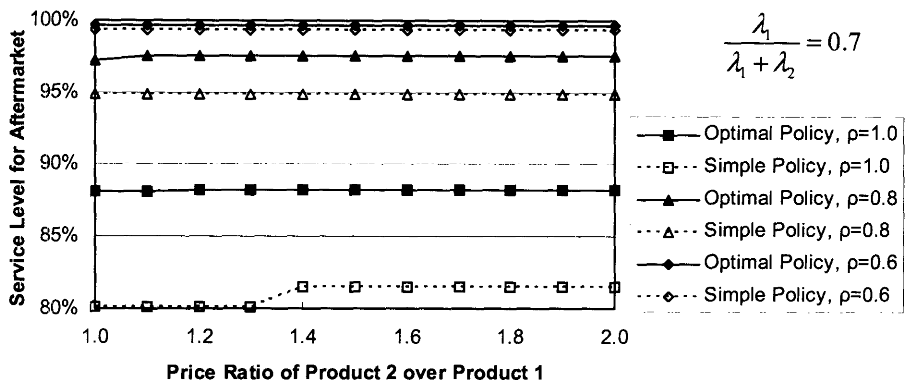


Figure 4-8: Service Level for Aftermarket Demands under the (S, R, B) Policy and the (S', B') Policy

Figure 4-9 illustrates service level for OEM demands. When the price ratio increases, the service level for OEM demand customer under both the optimal policy and the simple base-stock policy decreases. The figure also depicts that when the production capacity is sufficient, e.g., $\rho = 0.6, 0.8$, both policies maintain high service levels for OEM demands. As capacity becomes tight ($\rho = 1.0$), service level is lower under the optimal policy than the simple base-stock policy. This is because the optimal policy sacrifices the service level for OEM demands in order to accept more demands from the aftermarket and thus improve the overall profit. As ρ further increases, the difference in service level of the two policies becomes smaller and finally approaches zero.

We also compared the single-machine system with the two-machine system, and we observed that the optimal policy provides much higher service levels for aftermarket demands than the two-machine model, but lower service levels for OEM demands.

In summary, the optimal policy increases the expected profit as well as the fraction of satisfied aftermarket demands. This is particularly true when potential machine utilization is high.

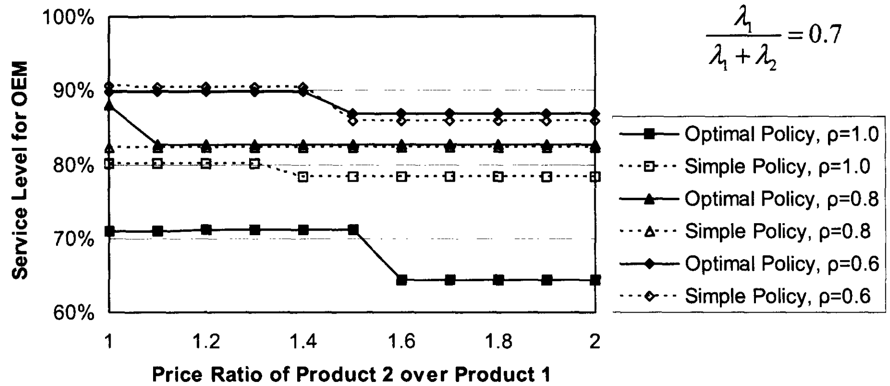


Figure 4-9: Service Level for OEM Demands under the (S, R, B) Policy and the (S', B') Policy

4.5 Concluding Remarks

In this chapter, we studied the production and inventory policies for two different make-to-stock/make-to-order manufacturing environments. For the model with both OEM and aftermarket demands, we show that the optimal policy can be characterized by three parameters, the base-stock level S, the rationing level R, and the admission level B. We extend the optimal policy to problems with cancelable aftermarket backorders by the manufacturer, as well as to problems with multiple types of aftermarket products. For the model where the manufacturer offers both customized and pre-configured (promotional) products, we show that the optimal policy can be characterized by two parameters, the basestock level S, and the admission level B. We also extend the optimal policy to problems with cancelable pre-configured backorders by the manufacturer.

After conducting an extensive computational analysis, we observed the following:

1. Implementing the optimal policy can result in up to 40% more profit than that under the base-stock policy. This difference in profit is high when production capacities is tight, when the backlogging penalty rates of the two products are close, or when demand arrival rates in the two classes are close.

2. When the manufacturer has the option of rejecting low priority orders, having the flexibility of canceling an accepted low-priority order does not provide a significant increase in profit.

3. Pooling production capacity between two types of products has the potential to improve the system profit dramatically. The pooling effect becomes more prominent when production capacities are tight, when the backlogging penalty rates of the two products are close, or when demand arrival rates in the two classes are close.

4. The optimal policy improves the profit by allocating the production capacity more effectively, and therefore having the capability to accept more low-priority demands. As a tradeoff, when the production capacity is very tight, the service level for the high-priority demands under the optimal policy is lower than that under the simple base-stock policy.

4.6 Appendix A

PROOF OF PROPOSITION 4.1

Proof. By contradiction, assume $R > S$. By Condition **C.3.2**, $\Delta_{12} f(S, -1) \geq 0$. Notice that $\Delta_{12} f(S, -1) = f(S+1, -1) - f(S, 0) + f(S+1, 0) - f(S+1, 0) = \Delta_1 f(S, 0) - \Delta_2 f(S+1, -1)$, so we get $\Delta_1 f(S, 0) \geq \Delta_2 f(S+1, -1)$, which cannot be true since we know that $\Delta_1 f(S, 0) < 0$ by Condition **C.2.2** and $\Delta_2 f(S+1, -1) \geq 0$ by Condition **C.1.1**. ■

PROOF OF LEMMA 4.1

We prove that function H preserves each condition in $\mathcal{C}$.

Proof for Condition C.1.1: We prove that $H_k f$ preserves Condition **C.1.1**, i.e., $\Delta_i f(\mathbf{y}) \geq 0$, for $y_i < 0$, and $i = 1, 2$. Since

$$\Delta_i H f(\mathbf{y}) = \Delta_i c(\mathbf{y}) + \Delta_i H_0 f(\mathbf{y}) + \Delta_i H_1 f(\mathbf{y}) + \Delta_i H_2 f(\mathbf{y}),$$

in order to prove that $\Delta_i H f(\mathbf{y}) \geq 0$ for $y_i < 0$, we show that all terms on the right hand side of the above equation are non-negative for $y_i < 0$ and $i = 1, 2$.

- For $H_0 f$, note that $H_0 f(\mathbf{y} + \mathbf{e_i})$ can be written as $H_0 f(\mathbf{y} + \mathbf{e_i}) = f(\mathbf{y} + \mathbf{e_i} + \mathbf{e_p})$, where $\mathbf{e_p}$ is one of the following three cases: $\mathbf{e_p} = \mathbf{e_0} = (0, 0)$, $\mathbf{e_p} = \mathbf{e_1} = (1, 0)$, or $\mathbf{e_p} = \mathbf{e_2} = (0, 1)$. Similarly, $H_0 f(\mathbf{y}) = f(\mathbf{y} + \mathbf{e_q})$, where $\mathbf{e_q}$ can be $\mathbf{e_0} = (0, 0)$, $\mathbf{e_1} = (1, 0)$, or $\mathbf{e_2} = (0, 1)$. Then we have

$$\begin{aligned}
\Delta_i H_0 f(\mathbf{y}) &= H_0 f(\mathbf{y} + \mathbf{e_i}) - H_0 f(\mathbf{y}) && (\mathit{Def.}) \\
&= f(\mathbf{y} + \mathbf{e_i} + \mathbf{e_p}) - f(\mathbf{y} + \mathbf{e_q}) && (\mathit{Above}) \\
&\geq f(\mathbf{y} + \mathbf{e_i} + \mathbf{e_q}) - f(\mathbf{y} + \mathbf{e_q}) && (\mathit{Max}) \\
&= \Delta_i f(\mathbf{y} + \mathbf{e_q}) && (\mathit{Def.}) \\
&\geq 0 && (\mathbf{C.1.1})
\end{aligned}$$

- For $H_1 f$, we have $H_1 f(\mathbf{y} + \mathbf{e_i}) = f(x + \mathbf{e_i} - \mathbf{e_1}) + p_1$ and $H_1 f(\mathbf{y}) = f(\mathbf{y} - \mathbf{e_1}) + p_1$. Then we will have

$$\Delta_i H_1 f(\mathbf{y}) = H_1 f(\mathbf{y} + \mathbf{e_i}) - H_1 f(\mathbf{y}) = f(\mathbf{y} + \mathbf{e_i} - \mathbf{e_1}) + p_1 - f(\mathbf{y} - \mathbf{e_1}) - p_1 = \Delta_i f(\mathbf{y} - \mathbf{e_1}) \geq 0.$$

- For H_2f, using a similar approach as in the first case, we can write $H_2f(\mathbf{y}+\mathbf{e_i}) = f(\mathbf{y}+\mathbf{e_i}-\mathbf{e_p})+g_2(p)$, and $H_2f(\mathbf{y}) = f(\mathbf{y}-\mathbf{e_q})+g_2(q)$ with $\mathbf{e_p}$ and $\mathbf{e_q}$ equal to either $\mathbf{e_0}$ or $\mathbf{e_2}$, and $g_2(0) = -r_2$ and $g_2(2) = p_2$. We have

$$
\begin{aligned}
\Delta_i H_2 f(\mathbf{y}) &= H_2 f(\mathbf{y}+\mathbf{e_i}) - H_2 f(\mathbf{y}) && (Def.)\\
&= f(\mathbf{y}+\mathbf{e_i}-\mathbf{e_p}) + g_2(p) - f(\mathbf{y}-\mathbf{e_q}) - g_2(q) && (Above)\\
&\geq f(\mathbf{y}+\mathbf{e_i}-\mathbf{e_q}) + g_2(q) - f(\mathbf{y}-\mathbf{e_q}) - g_2(q) && (Max)\\
&= \Delta_i f(\mathbf{y}-\mathbf{e_q}) && (Def.)\\
&\geq 0 && (\mathbf{C.1.1})
\end{aligned}
$$

Thus, we have $\Delta_i H_k f(\mathbf{y}) \geq 0$ for $k = 0, 1, 2$. On the other hand,

$$
\Delta_i c(\mathbf{y}) = c(\mathbf{y}+\mathbf{e_i}) - c(\mathbf{y}) = \begin{cases} b_1(y_1+1) + b_2y_2 - b_1y_1 - b_2y_2 = b_1 \geq 0 & if \quad y_1 < 0,\ y_2 \leq 0,\ i = 1\\ b_1y_1 + b_2(y_2+1) - b_1y_1 - b_2y_2 = b_2 \geq 0 & if \quad y_1 < 0,\ y_2 < 0,\ i = 2\\ -hy_1 + b_2(y_2+1) + hy_1 - b_2y_2 = b_2 \geq 0 & if \quad y_1 \geq 0,\ y_2 < 0,\ i = 2. \end{cases}
$$

so $\Delta_i c(\mathbf{y}) = b_i \geq 0$ if $y_i < 0$. Therefore, we have $\Delta_i Hf(\mathbf{y}) = \Delta_i c(\mathbf{y}) + \Delta_i H_0 f(\mathbf{y}) + \Delta_i H_1 f(\mathbf{y}) + \Delta_i H_2 f(\mathbf{y}) \geq 0$. This completes the proof for Condition **C.1.1**.

Proof for Condition C.1.2: The proof is similar to the proof for Condition **C.1.1**, and is therefore omitted.

<u>Proof for Condition C.2.1:</u> We must prove that $\Delta_1 H_k f(\mathbf{y})$, $\Delta_1 c(\mathbf{y})$, $\Delta_2 H_k f(\mathbf{y})$, and $\Delta_2 c(\mathbf{y})$ are non-increasing in y_1 and y_2.

<u>Proof for $\Delta_1 H_0 f(\mathbf{y})$:</u> First we prove that $\Delta_1 H_0 f(\mathbf{y})$ is non-increasing in y_1 and y_2. By definition $\Delta_1 H_0 f(\mathbf{y}) = H_0 f(\mathbf{y}+\mathbf{e_1}) - H_0 f(\mathbf{y})$, so we have,

$$\Delta_1 H_0 f(\mathbf{y}) = \begin{cases} f(\mathbf{y}+2\mathbf{e_1}) - f(\mathbf{y}+\mathbf{e_1}) = \Delta_1 f(\mathbf{y}+\mathbf{e_1}) & if \quad y_1 < R-1 \\ f(\mathbf{y}+2\mathbf{e_1}) - f(\mathbf{y}+\mathbf{e_1}) = \Delta_1 f(\mathbf{y}+\mathbf{e_1}) & if \quad y_1 = R-1, \quad y_2 = 0 \\ f(\mathbf{y}+\mathbf{e_1}+\mathbf{e_2}) - f(\mathbf{y}+\mathbf{e_1}) = \Delta_2 f(\mathbf{y}+\mathbf{e_1}) & if \quad y_1 = R-1, \quad y_2 < 0 \\ f(\mathbf{y}+2\mathbf{e_1}) - f(\mathbf{y}+\mathbf{e_1}) = \Delta_1 f(\mathbf{y}+\mathbf{e_1}) & if \quad R \le y_1 < S-1, \quad y_2 = 0 \\ f(\mathbf{y}+\mathbf{e_1}+\mathbf{e_2}) - f(\mathbf{y}+\mathbf{e_2}) = \Delta_1 f(\mathbf{y}+\mathbf{e_2}) & if \quad R \le y_1 < S-1, \quad y_2 < 0 \\ f(\mathbf{y}+\mathbf{e_1}) - f(\mathbf{y}+\mathbf{e_1}) = 0 & if \quad y_1 = S-1, \quad y_2 = 0 \\ f(\mathbf{y}+\mathbf{e_1}+\mathbf{e_2}) - f(\mathbf{y}+\mathbf{e_2}) = \Delta_1 f(\mathbf{y}+\mathbf{e_2}) & if \quad y_1 = S-1, \quad y_2 < 0 \\ f(\mathbf{y}+\mathbf{e_1}) - f(\mathbf{y}) = \Delta_1 f(\mathbf{y}) & if \quad y_1 \ge S, \quad y_2 = 0 \\ f(\mathbf{y}+\mathbf{e_1}+\mathbf{e_2}) - f(\mathbf{y}+\mathbf{e_2}) = \Delta_1 f(\mathbf{y}+\mathbf{e_2}) & if \quad y_1 \ge S, \quad y_2 < 0 \end{cases}$$

By **C.2.1**, $\Delta_1 f(\mathbf{y})$ and $\Delta_2 f(\mathbf{y})$ are both non-increasing in y_1 and y_2, so $\Delta_1 H_0(\mathbf{y})$ is non-increasing in y_1 and y_2 *within* each of the nine sub-condition intervals listed above. In the following, we will show that $\Delta_1 H_0(\mathbf{y})$ is non-increasing in y_1 and y_2 *across* any two adjacent intervals.

Here we only show that To show $\Delta_1 H_0 f(\mathbf{y})$ is non-increasing in y_1 and y_2 across the second and the third intervals (when $y_1 = R-1$ and $y_2 = 0$). The proofs for other intervals are similar and are therefore omitted.

let us denote $\mathbf{y^R} = (R-1, 0)$. By the results above, we have

$$\begin{aligned} \Delta_1 H_0 f(\mathbf{y^R}+\mathbf{e_1}) &= \Delta_1 f(\mathbf{y^R}+2\mathbf{e_1}) \\ \Delta_1 H_0 f(\mathbf{y^R}) &= \Delta_1 f(\mathbf{y^R}+\mathbf{e_1}) \\ \Delta_1 H_0 f(\mathbf{y^R}-\mathbf{e_1}) &= \Delta_1 f(\mathbf{y^R}). \end{aligned}$$

On the other hand, by **C.2.1**, we have

$$\Delta_1 f(\mathbf{y^R}+2\mathbf{e_1}) \le \Delta_1 f(\mathbf{y^R}+\mathbf{e_1}) \le \Delta_1 f(\mathbf{y^R}),$$

so $\Delta_1 H_0 f(\mathbf{y})$ is non-increasing in y_1 across the adjacent intervals at $\mathbf{y^R}$. To show that $\Delta_1 H_0 f(\mathbf{y})$ is also non-increasing in y_2 across the adjacent intervals at $\mathbf{y^R}$, note that

$$\Delta_1 H_0 f(\mathbf{y^R}-\mathbf{e_2}) = \Delta_2 f(\mathbf{y^R}-\mathbf{e_2}+\mathbf{e_1}) = \Delta_2 f(\mathbf{y^R}) \qquad \textbf{(C.4.1)}$$

On the other hand, by **C.2.1** and **C.3.2**, we have

$$\Delta_2 f(\mathbf{y^R}) \geq \Delta_2 f(\mathbf{y^R} + \mathbf{e_1}) \geq \Delta_1 f(\mathbf{y^R} + \mathbf{e_1}),$$

so $\Delta_1 H_0 f(\mathbf{y})$ is non-increasing in y_2 across the adjacent intervals at $\mathbf{y^R}$.

Proof for $\Delta_1 H_1 f(\mathbf{y})$: Now let us prove that $\Delta_1 H_1 f(\mathbf{y})$ is non-increasing in y_1 and y_2. We know,

$$\Delta_1 H_1 f(\mathbf{y}) = f(\mathbf{y} + \mathbf{e_1} - \mathbf{e_1}) + p_1 - f(\mathbf{y} - \mathbf{e_1}) - p_1 = \Delta_1 f(\mathbf{y} - \mathbf{e_1}).$$

By Condition **C.2.1**, $\Delta_1 f(\mathbf{y})$ is non-increasing in y_1 and y_2. So $\Delta_1 H_1 f(\mathbf{y})$ is also non-increasing in y_1 and y_2.

Proof for $\Delta_1 H_2 f(\mathbf{y})$: Now we prove that $\Delta_1 H_2 f(\mathbf{y})$ is non-increasing in y_1 and y_2. For this case, we have

$$\Delta_1 H_2 f(\mathbf{y}) = \begin{cases} f(\mathbf{y} + \mathbf{e_1} - \mathbf{e_2}) + p_2 - f(\mathbf{y} - \mathbf{e_2}) - p_2 = \Delta_1 f(\mathbf{y} - \mathbf{e_2}) & if \quad y_1 + y_2 > B \\ f(\mathbf{y} + \mathbf{e_1} - \mathbf{e_2}) + p_2 - f(\mathbf{y}) + r_2 \qquad = \Delta_{12} f(\mathbf{y} - \mathbf{e_2}) + p_2 + r_2 & if \quad y_1 + y_2 = B \\ f(\mathbf{y} + \mathbf{e_1}) - r_2 - f(\mathbf{y}) + r_2 \qquad = \Delta_1 f(\mathbf{y}) & if \quad y_1 + y_2 < B \end{cases}$$

Using Condition **C.2.1**, it is easy to show that $\Delta_1 H_2 f(\mathbf{y})$ is non-increasing in y_1 and y_2 when $y_1 + y_2 < B$ or when $y_1 + y_2 > B$. However, when $y_1 + y_2 = B$ the proof is different. Let us denote $\mathbf{y^B} = (y_1, y_2)$, such that $y_1 + y_2 = B$. In the following, we will compare $\Delta_1 H_2 f(\mathbf{y^B})$ with $\Delta_1 H_2 f(\mathbf{y^B} + \mathbf{e_1})$, $\Delta_1 H_2 f(\mathbf{y^B} + \mathbf{e_2})$, $\Delta_1 H_2 f(\mathbf{y^B} - \mathbf{e_1})$ and $\Delta_1 H_2 f(\mathbf{y^B} - \mathbf{e_2})$, one by one.

$$\begin{aligned}
\Delta_1 H_2 f(\mathbf{y^B}) - \Delta_1 H_2 f(\mathbf{y^B} + \mathbf{e_1}) &= \Delta_{12} f(\mathbf{y^B} - \mathbf{e_2}) + p_2 + r_2 - \Delta_1 f(\mathbf{y^B} - \mathbf{e_2} + \mathbf{e_1}) && (Above) \\
&= \Delta_{12} f(\mathbf{y^B}) + p_2 + r_2 - \Delta_1 f(\mathbf{y^B} - \mathbf{e_2} + \mathbf{e_1}) && (\mathbf{C.3.1}) \\
&\geq \Delta_1 f(\mathbf{y^B}) - \Delta_1 f(\mathbf{y^B} - \mathbf{e_2} + \mathbf{e_1}) && (\mathbf{C.4.2}) \\
&= f(\mathbf{y^B} + \mathbf{e_1}) - f(\mathbf{y^B}) - f(\mathbf{y^B} - \mathbf{e_2} + 2\mathbf{e_1}) + f(\mathbf{y^B} - \mathbf{e_2} + \mathbf{e_1}) \\
&= \Delta_{12} f(\mathbf{y^B} - \mathbf{e_2}) - \Delta_{12} f(\mathbf{y^B} - \mathbf{e_2} + \mathbf{e_1}) && (Def.) \\
&\geq 0 && (\mathbf{C.3.1})
\end{aligned}$$

$$\begin{aligned}
\Delta_1 H_2 f(\mathbf{y^B}) - \Delta_1 H_2 f(\mathbf{y^B} - \mathbf{e_1}) &= \Delta_{12} f(\mathbf{y^B} - \mathbf{e_2}) + p_2 + r_2 - \Delta_1 f(\mathbf{y^B} - \mathbf{e_1}) && (Above) \\
&\leq \Delta_1 f(\mathbf{y^B} - \mathbf{e_2}) - \Delta_1 f(\mathbf{y^B} - \mathbf{e_1}) && (\mathbf{C.4.2}) \\
&= f(\mathbf{y^B} - \mathbf{e_2} + \mathbf{e_1}) - f(\mathbf{y^B} - \mathbf{e_2}) - f(\mathbf{y^B}) + f(\mathbf{y^B} - \mathbf{e_1})
\end{aligned}$$

$$
\begin{aligned}
&= \Delta_{12}f(\mathbf{y^B}-\mathbf{e_2})-\Delta_{12}f(\mathbf{y^B}-\mathbf{e_2}-\mathbf{e_1}) && (\mathit{Def.})\\
&\leq 0 && \mathbf{(C.3.1)}
\end{aligned}
$$

Thus, $\Delta_1 H_2 f(\mathbf{y^B})$ is non-increasing in y_1 across the adjacent intervals at $\mathbf{y^B}$. We now compare $\Delta_1 H_2 f(\mathbf{y^B})$ with $\Delta_1 H_2 f(\mathbf{y^B}+\mathbf{e_2})$ and $\Delta_1 H_2 f(\mathbf{y^B}-\mathbf{e_2})$ to show that $\Delta_1 H_2 f(\mathbf{y^B})$ is also non-increasing in y_2 across the adjacent intervals.

$$
\begin{aligned}
\Delta_1 H_2 f(\mathbf{y^B})-\Delta_1 H_2 f(\mathbf{y^B}+\mathbf{e_2}) &= \Delta_{12}f(\mathbf{y^B}-\mathbf{e_2})+p_2+r_2-\Delta_1 f(\mathbf{y^B}) && (\mathit{Above})\\
&= \Delta_{12}f(\mathbf{y^B})+p_2+r_2-\Delta_1 f(\mathbf{y^B}) && \mathbf{(C.3.1)}\\
&= p_2+r_2-\Delta_2 f(\mathbf{y^B}) && (\mathit{Def.})\\
&\geq 0 && \mathbf{(C.4.2)}
\end{aligned}
$$

$$
\begin{aligned}
\Delta_1 H_2 f(\mathbf{y^B})-\Delta_1 H_2 f(\mathbf{y^B}-\mathbf{e_2}) &= \Delta_{12}f(\mathbf{y^B}-\mathbf{e_2})+p_2+r_2-\Delta_1 f(\mathbf{y^B}-\mathbf{e_2}) && (\mathit{Above})\\
&= p_2+r_2-\Delta_2 f(\mathbf{y^B}-\mathbf{e_2}) && (\mathit{Def.})\\
&\leq 0 && \mathbf{(C.4.2)}
\end{aligned}
$$

Therefore, $\Delta_1 H_2 f(\mathbf{y^B})$ is also non-increasing in y_2 across the adjacent intervals at $\mathbf{y^B}$.

Proof for $\Delta_1 c(\mathbf{y})$: We now show that $\Delta_1 c(\mathbf{y})$ is non-increasing in y_1, y_2. Note that,

$$
\Delta_1 c(\mathbf{y}) = c(\mathbf{y}+\mathbf{e_1})-c(\mathbf{y}) = \begin{cases} b_1(y_1+1)+b_2y_2-b_1y_1-b_2y_2 = b_1 \geq 0 & if \quad y_1<0\\ -h(y_1+1)+b_2y_2+hy_1-b_2y_2=-h\leq 0 & if \quad y_1\geq 0 \end{cases}
$$

which is is non-increasing in y_1, y_2.

In conclusion, since $\Delta_1 H_0 f(\mathbf{y})$, $\Delta_1 H_1 f(\mathbf{y})$, $\Delta_1 H_2 f(\mathbf{y})$, and $\Delta_1 c(\mathbf{y})$ are all non-increasing in y_1, y_2, then $\Delta_1 H f(\mathbf{y})$ is non-increasing in y_1, y_2. This completes the proof for $\Delta_1 H f(\mathbf{y})$.

Proof for $\Delta_2 H_0 f(\mathbf{y})$: Next, we prove that $\Delta_2 H_k f(\mathbf{y})$ is non-increasing in y_1 and y_2. By definition, $\Delta_2 H_k f(\mathbf{y}) = H_k f(\mathbf{y}+\mathbf{e_2})-H_k f(\mathbf{y})$, so we have,

$$
\Delta_2 H_0 f(\mathbf{y}) = \begin{cases} f(\mathbf{y}+\mathbf{e_2}+\mathbf{e_1})-f(\mathbf{y}+\mathbf{e_1}) = \Delta_2 f(\mathbf{y}+\mathbf{e_1}) & if \quad y_1<R\\ f(\mathbf{y}+2\mathbf{e_2})-f(\mathbf{y}+\mathbf{e_2}) \qquad = \Delta_2 f(\mathbf{y}+\mathbf{e_2}) & if \quad y_1\geq R, \quad y_2<-1\\ f(\mathbf{y}+\mathbf{e_2}+\mathbf{e_1})-f(\mathbf{y}+\mathbf{e_2}) = \Delta_1 f(\mathbf{y}+\mathbf{e_2}) & if \quad S>y_1\geq R, \quad y_2=-1\\ f(\mathbf{y}+\mathbf{e_2})-f(\mathbf{y}+\mathbf{e_2}) \qquad = 0 & if \quad y_1\geq S, \quad y_2=-1 \end{cases}
$$

Using the same argument as in the proof for $\Delta_1 H_0 f(\mathbf{y})$, it is easy to show that $\Delta_2 H_0(\mathbf{y})$ is non-increasing in y_1 and y_2.

Proof for $\Delta_2 H_1 f(\mathbf{y})$: Now we prove that $\Delta_2 H_1 f(\mathbf{y})$ is non-increasing in y_1 and y_2, we have

$$\Delta_2 H_1 f(\mathbf{y}) = f(\mathbf{y} + \mathbf{e_2} - \mathbf{e_1}) + p_1 - f(\mathbf{y} - \mathbf{e_1}) - p_1 = \Delta_2 f(\mathbf{y} - \mathbf{e_1}).$$

By Condition **C.2.1**, we know that $\Delta_2 f(\mathbf{y})$ is non-increasing in y_1 and y_2. Thus, $\Delta_2 H_1 f(\mathbf{y})$ is also non-increasing in y_1 and y_2.

Proof for $\Delta_2 H_2 f(\mathbf{y})$: Now we prove that $\Delta_2 H_2 f(\mathbf{y})$ is non-increasing in y_1 and y_2, we have

$$\Delta_2 H_2 f(\mathbf{y}) = \begin{cases} f(\mathbf{y} + \mathbf{e_2} - \mathbf{e_2}) + p_2 - f(\mathbf{y} - \mathbf{e_2}) - p_2 = \Delta_2 f(\mathbf{y} - \mathbf{e_2}) & if \quad y_1 + y_2 > B \\ f(\mathbf{y} + \mathbf{e_2} - \mathbf{e_2}) + p_2 - f(\mathbf{y}) + r_2 \qquad = p_2 + r_2 & if \quad y_1 + y_2 = B \\ f(\mathbf{y} + \mathbf{e_2}) - r_2 - f(\mathbf{y}) + r_2 \qquad = \Delta_2 f(\mathbf{y}) & if \quad y_1 + y_2 < B \end{cases}$$

According to Condition **C.2.1**, we know that $\Delta_1 f(\mathbf{y})$ and $\Delta_2 f(\mathbf{y})$ are non-increasing in y_1 and y_2. Therefore, $\Delta_2 H_2(\mathbf{y})$ is non-increasing in y_1 and y_2 when $y_1 + y_2 < B$ or when $y_1 + y_2 > B$. Using the same argument as in the proof of $\Delta_1 H_2 f(\mathbf{y})$, it is easy to prove that $\Delta_2 H_2(\mathbf{y})$ is non-increasing in y_1 and y_2.

Proof for $\Delta_2 c(\mathbf{y})$: Finally, note that $\Delta_2 c(\mathbf{y}) = c(\mathbf{y} + \mathbf{e_2}) - c(\mathbf{y}) = b_2$ is independent of y_1 and y_2, and therefore non-increasing in y_1 and y_2. Since $\Delta_2 H_0 f(\mathbf{y})$, $\Delta_2 H_1 f(\mathbf{y})$, $\Delta_2 H_2 f(\mathbf{y})$, and $\Delta_2 c(\mathbf{y})$ are all non-increasing in y_1, y_2, then $\Delta_2 H f(\mathbf{y})$ is non-increasing in y_1, y_2. This completes the proof for $\Delta_2 H f(\mathbf{y})$, and therefore the proof for Condition **C.2.1**.

Proof for Condition C.2.2: Condition **C.2.2** is a direct result of Condition **C.2.1**.

Proof for Condition C.3.1: We prove that $\Delta_{12} H_k f(\mathbf{y})$ and $\Delta_{12} c(\mathbf{y})$ are non-increasing in y_1 and is independent of y_2. By definition, $\Delta_{12} H_k f(\mathbf{y}) = H_k f(\mathbf{y} + \mathbf{e_1}) - H_k f(\mathbf{y} + \mathbf{e_2})$, so we have,

$$\Delta_{12} H_0 f(\mathbf{y}) \;=\; \begin{cases} f(\mathbf{y} + \mathbf{e_1} + \mathbf{e_2}) - f(\mathbf{y} + 2\mathbf{e_2}) \qquad = \Delta_{12} f(\mathbf{y} + \mathbf{e_2}) & if \quad y_1 > R - 1 \\ f(\mathbf{y} + \mathbf{e_1} + \mathbf{e_2}) - f(\mathbf{y} + \mathbf{e_2} + \mathbf{e_1}) = 0 & if \quad y_1 = R - 1 \\ f(\mathbf{y} + \mathbf{e_1} + \mathbf{e_1}) - f(\mathbf{y} + \mathbf{e_2} + \mathbf{e_1}) = \Delta_{12} f(\mathbf{y} + \mathbf{e_1}) & if \quad y_1 < R - 1 \end{cases}$$

$$\Delta_{12} H_1 f(\mathbf{y}) \;=\; f(\mathbf{y} + \mathbf{e_1} - \mathbf{e_1}) + p_1 - f(\mathbf{y} + \mathbf{e_2} - \mathbf{e_1}) - p_1 = \Delta_{12} f(\mathbf{y} - \mathbf{e_1})$$

$$\Delta_{12}H_2f(\mathbf{y}) = \begin{cases} f(\mathbf{y}+\mathbf{e_1}-\mathbf{e_2})+p_2-f(\mathbf{y}+\mathbf{e_2}-\mathbf{e_2})-p_2 = \Delta_{12}f(\mathbf{y}-\mathbf{e_2}) & if \quad y_1+y_2 \geq B \\ f(\mathbf{y}+\mathbf{e_1})-r_2-f(\mathbf{y}+\mathbf{e_2})+r_2 \qquad\qquad\qquad = \Delta_{12}f(\mathbf{y}) & if \quad y_1+y_2 < B \end{cases}$$

Using the same argument as in the proof for Condition **C.2.1**, it is easy to show that $\Delta_{12}H_kf(\mathbf{y})$ is non-increasing in y_1 and independent in y_2.

Note that,

$$\Delta_{12}c(\mathbf{y}) = \begin{cases} b_1 - b_2 & if \quad y_1 < 0 \\ -h_1 - b_2 & if \quad y_1 \geq 0 \end{cases}$$

which is also non-increasing in y_1 and independent in y_2. In conclusion, $\Delta_{12}Hf(\mathbf{y})$ is non-increasing in y_1 and is independent in y_2.

Proof for Condition C.3.2: Condition **C.3.2** is a direct result of Condition **C.3.1**.

Proof for Condition C.4.1: Condition **C.4.1** can be derived from Condition **C.3.1**. Since $\Delta_{12}f(y_1, y_2)$ is independent in y_2, we have

$$\begin{array}{rrcll} & \Delta_{12}f(y_1-1,y_2) & = & \Delta_{12}f(y_1-1,y_2-1) & (\mathbf{C.3.1}) \\ \Rightarrow & f(y_1,y_2)-f(y_1-1,y_2+1) & = & f(y_1,y_2-1)-f(y_1-1,y_2) & (Def.) \\ \Rightarrow & f(y_1,y_2)-f(y_1,y_2-1) & = & f(y_1-1,y_2+1)-f(y_1-1,y_2) & (Rearrangement) \\ \Rightarrow & \Delta_2f(y_1,y_2-1) & = & \Delta_2f(y_1-1,y_2) & (Def.) \end{array}$$

Thus, in general we have $\Delta_2f(y_1,y_2) = \Delta_2f(y_1 \mp 1, y_2 \pm 1)$, which directly implies Condition **C.4.1** (Note that Condition **C.4.1** can be obtained by adding 1 to one element and deducing 1 form the other element of vector $\mathbf{y}$ in the above result.

Proof for Condition C.4.2: Condition **C.4.2** is a direct result of Condition **C.4.1**. ∎

PROOF OF THEOREM 4.1

We first prove the existence of an optimal stationary policy by following the approaches in [17] and [18]. For this purpose, we need to show: (i) The set of structured functions $\mathcal{C}$ is complete, and (ii) $J(y_1, y_2) \in \mathcal{C}$.

By Lemma 4.1, the composite operator $H = H_0 \circ H_1 \circ H_2$ preserves the structural properties from Condition **C.1.1** to **C.4.2**. Because the limit of any converging sequences of functions in $\mathcal{C}$ will be in $\mathcal{C}$ as well, the set of structured functions $\mathcal{C}$ is complete. On the other hand, since $c(y_1, y_2) \in \mathcal{C}$, Lemma 1 implies that $J(y_1, y_2) \in \mathcal{C}$. So the optimal expected profit function J is structured and satisfies all conditions in $\mathcal{C}$. Hence, the existence of an optimal stationary policy under the discounted-profit criterion follows from the corollaries of Theorem 5.1 of [26].

To argue the structural properties are retained for the average-profit case, we use the three conditions in Sennott (1999, p. 132). Let $(0,0)$ be the distinguished state, it is easy to show that SEN1, SEN2 and SEN3 are satisfied in our MDP model. Therefore, the average-profit problem can be obtained as the limit of discounted-profit problem by letting $\alpha \to 0^+$.

From equation (4.3), we can see that when there is no backorder for any product, the optimal production control only depends on the sign of $\Delta_1 J(y_1, y_2)$: to produce if the sign is positive, and to stop otherwise. By Condition **C.2.2**, it is optimal to produce if $y_1 < S$.

When there are backorders for both products, by Condition **C.1.1**, it is always optimal to produce, and by Condition **C.1.2**, the manufacturer will always give priority to type 1. If there are only backorders for type 2, by Conditions **C.3.1** and **C.3.2**, the manufacturer will produce type 2 if $y_1 \geq R$, or otherwise produce type 1 to build up the inventory.

The admission control for an arriving order for type 2 depends on the sign of $p_2 + r_2 - \Delta_2 J(y_1, y_2 - 1)$: satisfy or backlog the demand if the sign is positive, reject otherwise. So by Conditions **C.4.1** and **C.4.2**, the optimal routing policy is to backlog an arriving type 2 demand if $y_1 + y_2 > B$, and to reject otherwise. ∎

PROOF OF PROPOSITION 4.2

Proof. By contradiction, assume $L > B$. By Condition **C.4.2**, $\Delta_2 f(0, L-1) \leq p_2 + r_2 < p_2 + l_2$. This cannot be true, since by Condition **C.4.3**, we have $\Delta_1 f(0, L-1) > p_2 + l_2$. ∎

PROOF OF THEOREM 4.2

Proof. The proof of the production, rationing, and admission control policies are the same as in Theorem 1, so we only discuss the cancelation control policy.

When an order for type 1 arrives, the cancelation control for type 2 backorders depends on the sign of $p_2 + l_2 - \Delta_2 J(y_1, y_2)$: cancel a backorder if the sign is negative, and do not cancel otherwise. So by Conditions **C.2.1** and **C.4.3**, the optimal cancelation policy is to cancel a type 2 backorder if $y_1 + y_2 \leq L$. ∎

PROOF OF PROPOSITION 4.3

Proof. When $2 \leq i < j$, $p_i + r_i \leq p_j + r_j$, by Condition **C.2.1**, $\Delta_2 f(y_1, y_2)$ is non-increasing in y_2, so $B_i \geq B_j$. ∎

4.7 Appendix B

PROOF OF LEMMA 4.2

We show that function H preserves each condition in $\mathcal{D}$. Proofs for Conditions **D.1.1** and **D.1.2** are similar to the proofs for Conditions **C.1.1** and **C.1.2**, and are therefore omitted.

Proof for Condition D.2.1:

Proof for $\Delta_1 H_0 f(\mathbf{y})$: We first prove that $\Delta_1 H_0 f(\mathbf{y})$ is non-increasing in y_1 and y_2. By definition, $\Delta_1 H_0 f(\mathbf{y}) = H_0 f(\mathbf{y} + \mathbf{e_1}) - H_0 f(\mathbf{y})$, so we have,

$$\Delta_1 H_0 f(\mathbf{y}) = \begin{cases} f(\mathbf{y} + \mathbf{e_1} + \mathbf{e_2}) - f(\mathbf{y} + \mathbf{e_2}) = \Delta_1 f(\mathbf{y} + \mathbf{e_2}) & if \quad y_2 < 0 \\ f(\mathbf{y} + 2\mathbf{e_1}) - f(\mathbf{y} + \mathbf{e_1}) \qquad = \Delta_1 f(\mathbf{y} + \mathbf{e_1}) & if \quad y_2 = 0, \quad y_1 < S - 1 \\ f(\mathbf{y} + \mathbf{e_1}) - f(\mathbf{y} + \mathbf{e_1}) \qquad = 0 & if \quad y_2 = 0, \quad y_1 = S - 1 \\ f(\mathbf{y} + \mathbf{e_1}) - f(\mathbf{y}) \qquad\qquad = \Delta_1 f(\mathbf{y}) & if \quad y_2 = 0, \quad y_1 > S - 1 \end{cases}$$

By **D.2.1**, $\Delta_1 f(\mathbf{y})$ is non-increasing in y_1 and y_2, so $\Delta_1 H_0 f(\mathbf{y})$ is non-increasing in y_1 and y_2 *within* each of the above four sub-condition intervals. In the following, we show that $\Delta_1 H_0 f(\mathbf{y})$ is also non-increasing in y_1 and y_2 *across* any adjacent intervals. We only present the proof for the case where $y_1 = S - 1$ and $y_2 = 0$. The proof for other cases are similar and are therefore omitted.

When $y_1 = S-1$ and $y_2 = 0$, let us denote $\mathbf{y^S} = (S-1, 0)$. Then we have,

$$\begin{aligned}
\Delta_1 H_0 f(\mathbf{y^S} + \mathbf{e_1}) &= \Delta_1 f(\mathbf{y^S} + \mathbf{e_1}) \\
\Delta_1 H_0 f(\mathbf{y^S}) &= 0 \\
\Delta_1 H_0 f(\mathbf{y^S} - \mathbf{e_1}) &= \Delta_1 f(\mathbf{y^S})
\end{aligned}$$

By **D.2.2**, we have $\Delta_1 f(\mathbf{y^S} + \mathbf{e_1}) \leq 0 \leq \Delta_1 f(\mathbf{y^S})$, so $\Delta_1 H_0 f(\mathbf{y})$ is non-increasing in y_1 across the adjacent intervals at $\mathbf{y^S}$.

Notice that,

$$\Delta_1 H_0 f(\mathbf{y^S} - \mathbf{e_2}) = \Delta_1 f(\mathbf{y^S})$$

so for the same reason, $\Delta_1 H_0 f(\mathbf{y})$ is non-increasing in y_2 across the adjacent intervals at $\mathbf{y^S}$.

Proof for $\Delta_1 H_1 f(\mathbf{y})$: Now we prove that $\Delta_1 H_1 f(\mathbf{y})$ is non-increasing in y_1 and y_2. By definition, $\Delta_1 H_1 f(\mathbf{y}) = H_1 f(\mathbf{y} + \mathbf{e_1}) - H_1 f(\mathbf{y})$, so we have,

$$\Delta_1 H_1 f(\mathbf{y}) = \begin{cases}
f(\mathbf{y} + \mathbf{e_1} - \mathbf{e_1}) + p_1 - f(\mathbf{y} - \mathbf{e_1}) - p_1 = \Delta_1 f(\mathbf{y} - \mathbf{e_1}) & if \quad y_1 > 0 \\
f(\mathbf{y} + \mathbf{e_1} - \mathbf{e_1}) + p_1 - f(\mathbf{y} - \mathbf{e_1}) - p_1 = \Delta_1 f(\mathbf{y} - \mathbf{e_1}) & if \quad y_1 + y_2 > B, \quad y_1 \leq 0 \\
f(\mathbf{y} + \mathbf{e_1} - \mathbf{e_1}) + p_1 - f(\mathbf{y}) + r_1 \qquad = p_1 + r_1 & if \quad y_1 + y_2 = B, \quad y_1 \leq 0 \\
f(\mathbf{y} + \mathbf{e_1}) - r_1 - f(\mathbf{y} - \mathbf{e_1}) + r_1 \qquad = \Delta_1 f(\mathbf{y}) & if \quad y_1 + y_2 < B, \quad y_1 \leq 0
\end{cases}$$

By **D.2.1**, $\Delta_1 f(\mathbf{y})$ is non-increasing in y_1 and y_2, so $\Delta_1 H_1 f(\mathbf{y})$ is non-increasing in y_1 and y_2 *within* each of the four sub-condition intervals. In the following, we show that $\Delta_1 f(\mathbf{y})$ is non-increasing in y_1 and y_2 *across* any two adjacent intervals. We only present the proof for the third interval, since the proofs for other intervals are similar.

Let us denote $\mathbf{y^B} = (y_1, y_2)$, such that $y_1 + y_2 = B$ and $y_1 \leq 0$. With the above results, we have,

$$\begin{aligned}
\Delta_1 H_1 f(\mathbf{y^B} + \mathbf{e_1}) &= \Delta_1 f(\mathbf{y^B}) \\
\Delta_1 H_1 f(\mathbf{y^B}) &= p_1 + r_1 \\
\Delta_1 H_1 f(\mathbf{y^B} - \mathbf{e_1}) &= \Delta_1 f(\mathbf{y^B} - \mathbf{e_1}) \\
\Delta_1 H_1 f(\mathbf{y^B} + \mathbf{e_2}) &= \Delta_1 f(\mathbf{y^B} + \mathbf{e_2} - \mathbf{e_1}) \\
\Delta_1 H_1 f(\mathbf{y^B} - \mathbf{e_2}) &= \Delta_1 f(\mathbf{y^B} - \mathbf{e_2})
\end{aligned}$$

Condition **D.3.1** and **D.3.2** imply that $\Delta_1 f(\mathbf{y^B}+\mathbf{e_2}-\mathbf{e_1}) = \Delta_1 f(\mathbf{y^B}) \le p_1 + r_1 \le \Delta_1 f(\mathbf{y^B}-\mathbf{e_1}) = \Delta_1 f(\mathbf{y^B}-\mathbf{e_2})$, so $\Delta_1 H_1 f(\mathbf{y})$ is non-increasing in y_1 and y_2 across the adjacent intervals at $\mathbf{y^B}$.

Proof for $\Delta_1 H_2 f(\mathbf{y})$: Now we prove that $\Delta_1 H_2 f(\mathbf{y})$ is non-increasing in y_1 and y_2. We have

$$\Delta_1 H_2 f(\mathbf{y}) = H_2 f(\mathbf{y}+\mathbf{e_1}) - H_2 f(\mathbf{y}) = f(\mathbf{y}+\mathbf{e_1}-\mathbf{e_2}) - f(\mathbf{y}-\mathbf{e_2}) = \Delta_1 f(\mathbf{y}-\mathbf{e_2}).$$

By **D.2.1**, it is clear that $\Delta_1 H_2 f(\mathbf{y})$ is non-increasing in y_1 and y_2.

Proof for $\Delta_1 c(\mathbf{y})$: Finally for $c_1(\mathbf{y})$ we have

$$\Delta_1 c_1(\mathbf{y}) = c_1(\mathbf{y}+\mathbf{e_1}) - c_1(\mathbf{y}) = \begin{cases} b_1(y_1+1) - b_1 y_1 = b_1 & if \quad y_1 < 0; \\ -h(y_1+1) + h y_1 = -h & if \quad y_1 \ge 0. \end{cases}$$

which is non-increasing in y_1 and y_2. In Conclusion, $\Delta_1 H f(\mathbf{y})$ is non-increasing in y_1 and y_2.

Proof for Condition D.2.2: Condition **D.2.2** is a direct result of Condition **D.2.1**.

Proof for Condition D.3.1: We first prove that $\Delta_{12} f(y_1, y_2) = \Delta_{12} f(y_1 - 1, y_2)$ for $y_1 < 0$. By definition, $\Delta_{12} H_0 f(\mathbf{y}) = H_0 f(\mathbf{y}+\mathbf{e_1}) - H_0 f(\mathbf{y}+\mathbf{e_2})$, so we have,

$$\begin{aligned}
\Delta_{12} H_0 f(\mathbf{y}) &= \begin{cases} f(\mathbf{y}+\mathbf{e_1}+\mathbf{e_2}) - f(\mathbf{y}+2\mathbf{e_2}) & = \Delta_{12} f(\mathbf{y}+\mathbf{e_2}) \quad if \quad y_2 < -1 \\ f(\mathbf{y}+\mathbf{e_1}+\mathbf{e_2}) - f(\mathbf{y}+\mathbf{e_2}+\mathbf{e_1}) = 0 & \qquad\qquad\qquad\quad if \quad y_2 = -1 \end{cases} \\
\Delta_{12} H_1 f(\mathbf{y}) &= \begin{cases} f(\mathbf{y}+\mathbf{e_1}-\mathbf{e_1}) + p_1 - f(\mathbf{y}+\mathbf{e_2}-\mathbf{e_1}) - p_1 = \Delta_{12} f(\mathbf{y}-\mathbf{e_1}) & if \quad y_1 + y_2 \ge B \\ f(\mathbf{y}+\mathbf{e_1}) - r_1 - f(\mathbf{y}+\mathbf{e_2}) + r_1 \qquad\qquad = \Delta_{12} f(\mathbf{y}) & if \quad y_1 + y_2 < B \end{cases} \\
\Delta_{12} H_2 f(\mathbf{y}) &= \quad f(\mathbf{y}+\mathbf{e_1}-\mathbf{e_2}) - f(\mathbf{y}+\mathbf{e_2}-\mathbf{e_2}) = \Delta_{12} f(\mathbf{y}-\mathbf{e_2})
\end{aligned}$$

Noting that $\Delta_{12} c_1(\mathbf{y}) = b_1$ is independent of y_1 for $y_1 < 0$, and $\Delta_{12} f(y_1, y_2) = \Delta_{12} f(y_1 - 1, y_2)$, so $\Delta_{12} H f(y_1, y_2) = \Delta_{12} H f(y_1 - 1, y_2)$.

Then we have

$$\begin{aligned}
\Delta_{12} f(\mathbf{y}) &= \Delta_{12} f(\mathbf{y}+\mathbf{e_1}) \\
\Rightarrow \qquad f(\mathbf{y}+\mathbf{e_1}) - f(\mathbf{y}+\mathbf{e_2}) &= f(\mathbf{y}+\mathbf{e_1}+\mathbf{e_1}) - f(\mathbf{y}+\mathbf{e_1}+\mathbf{e_2}) \qquad (Def.)
\end{aligned}$$

$$\Rightarrow \quad f(\mathbf{y}+\mathbf{e_1}+\mathbf{e_2}) - f(\mathbf{y}+\mathbf{e_2}) = f(\mathbf{y}+\mathbf{e_1}+\mathbf{e_1}) - f(\mathbf{y}+\mathbf{e_1}) \qquad (Rearrangement)$$

$$\Rightarrow \quad \Delta_1 f(\mathbf{y}+\mathbf{e_2}) = \Delta_1 f(\mathbf{y}+\mathbf{e_1}) \qquad (Def.)$$

Therefore, we have $\Delta_1 f(\mathbf{y}) = \Delta_1 f(\mathbf{y}+\mathbf{e_1}-\mathbf{e_2})$, and Condition **D.3.1** can be derived by adding 1 to one element and deducing 1 form another element in vector $\mathbf{y}$ in the above result.

Proof for Condition D.3.2: Condition **D.3.2** can be directly deduced from Conditions **D.2.1** and **D.3.1**. ■

PROOF OF THEOREM 4.3

The proof for the existence of the optimal policy is similar to the proof in Theorem 4.1.

From equation 4.8, we see that when there are no backorders from any classes, the optimal production control only depends on the sign of $\Delta_1 J_1(y_1, y_2)$: produce if it is positive, or do not produce otherwise. So by Condition **D.2.2**, it is optimal to produce if $y_1 < S$. When there are backorders, by Conditions **D.1.1** and **D.1.2**, it is always optimal to produce for the backorder with the highest priority.

The admission control for an arriving demand for type 1 depends on the sign of $p_1 + r_1 - \Delta_1 J_1(y_1 - 1, y_2)$: satisfy or backlog the demand if the sign is positive; reject otherwise. So, by Conditions **D.2.1** and **D.3.2**, the optimal admission policy is to satisfy an arriving demand for type 1 from inventory if $y_1 > 0$, to backlog if $y_1 < 0$ and $y_1 + y_2 > B$, and to reject if $y_1 + y_2 \leq B$. ■

PROOF OF PROPOSITION 4.4

By contradiction, assume $L > B$. By Condition **D.3.2**, $\Delta_1 f(L-1, 0) \leq p_1 + r_1 < p_1 + l_1$. This cannot be true, since by Condition **D.3.3** we have $\Delta_1 f(L-1, 0) > p_1 + l_1$. ■

PROOF OF THEOREM 4.4

The proof of the production and admission control policies are the same as in Theorem 4.3, so we only discuss the cancelation control policy.

After accepting a demand for a higher-priority product, the cancelation control for type 1 backorders depends on the sign of $p_1 + l_1 - \Delta_1 J_1(y_1, y_2)$: cancel a backorder if the value is negative, and do not cancel otherwise. So by Conditions **D.2.2** and **D.3.2**, the optimal cancelation policy is to cancel a type 1 backorder if $y_1 + y_2 \leq L$. $\blacksquare$

Chapter 5

Conclusions and Future Directions

In this chapter, we summarize our main results and suggest future research directions.

In Chapter 2, we analyze a single-class-customer problem with both backlogged and discretionary sales. We analyze the structure of the optimal policy and show that it is characterized by three state-independent control parameters: the produce-up-to level (S), the reserve-up-to level (R), and the backlog-up-to level (B). In such a policy, at the beginning of each period, the manufacturer will produce to bring the inventory level up to S or to the maximum capacity; during the period, s/he will set aside R units of inventory for the next period, and satisfy some customers with the remaining inventory, if expected future profit is higher; otherwise, s/he will satisfy customers with the inventory and backlog up to B units of demands.

In Chapter 3, we study two single-product, two-class-customer models in which high-priority customers would not wait for delayed fulfillments, while other customers are willing to wait in order to pay lower prices. For the first model, where the manufacturer does not differentiate customers and serve them as a single class, we show that the structure of the optimal policy is characterized by three state-independent control parameters: the produce-up-to level (S), the reserve-up-to level (R), and the backlog-up-to level (B). In the second model, the manufacturer differentiates customers according to their sensitivities to leadtime, and charge higher prices to customers who want the product immediately. For this model, we provide a heuristic policy characterized by three threshold levels: S, R, B. In such a policy, during each period, the manufacturer will set aside R units of inventory for the next period, and satisfy some high priority customers with the remaining inventory, if expected future profit is higher; otherwise, s/he will satisfy as many of the high priority customers as possible and backlog up to B units of lower priority customers.

In Chapter 4, we examine manufacturing systems that produce both make-to-stock and make-to-order products. Two models are studied: In the first model, which is motivated by the practice in the automobile industry, the make-to-stock (OEM) product has higher priority than the make-to-order (aftermarket) product; In the second model, which is motivated by applications in the computer industry, the make-to-order (customized) product has higher priority than the make-to-stock (promotional) product. For both models, we analyze the structure of the optimal production and admission policies and show that it is characterized by linear threshold levels. We also extend those results to problems where low-priority backorders can be canceled by the manufacturer, as well as to problems with multiple types of make-to-order products.

Finally, we discuss some future research directions.

(1) **Joint pricing and allocation decisions with multiple products:** This dissertation raises an important future research direction, the joint pricing and capacity allocation decisions for multiple products. So far, academic literature on joint pricing and allocation decisions has confined itself mainly to single-product and single-class-customer problems, which are not applicable to the current practice of market segmentation in the manufacturing industry. The complexity of the joint pricing and allocation problem among multiple products increases dramatically due to demand diversions among different products, i.e., demand for one product not only depends on its own price, but may also be influenced by the prices of other products. Another issue complicating the problem is that different products may compete for the same resources, i.e., the production capacity and components.

(2) **Multiple manufacturers:** Competition is an important issue in the real world.

Generally speaking, both price and service (lead time) are important in the manufacturer's ability to gain a certain amount of market share. Of course, the impact of price and lead time on the manufacturer's ability to compete varies from industry to industry. For example, for medical equipment component suppliers, fast fulfillment is very important; while for computer component manufacturers, competition is focused on price. So it is crucial for manufacturers to understand customers' preference and make the effective pricing and capacity decisions.

(3) **Capacity expansion:** In this dissertation, we assume that the production capacity is fixed. In reality, the manufacturer may buy or lease some additional capacity when he/she finds it is necessary. In addition, the manufacturer may adjust the production capacity if demand has strong seasonality. Hence, the challenge is to incorporate the option of capacity expansion into the model.

(4) **Production lead time and other issues:** In this dissertation, we assume zero production lead time in periodic review models, or exponential production time in continuous review models. If production time is generally distributed, the state space has to be expanded, and thus the problem becomes more difficult to analyze. There are also other issues to be incorporated into the model, such as general demand distribution functions, production set-up time/cost for a single product, or switching times/costs between multiple products.

Bibliography

[1] Agrawal, V. and A. Kambil. 2000. Dynamic pricing strategies in electronic commerce. Working paper, Stern Business School, New York University.

[2] Baras, J.S., D.J.Ma and A.M.Makowski. 1985. K competing queues with geometric service requirements and linear costs: the μc rule is always optimal.*Systems Control Letters*, Vol. 6, pp. 173-180.

[3] Bertsekas, D. 1995. *Dynamic Programming and Optimal Control*, Volume One, Athena Scientific.

[4] Buyukkoc, C., P.Varaiya and J. Walrand. 1985. The $c\mu-$rule revisited. *Advances in Applied Probability*, Vol. 17, pp. 237-238.

[5] Carr, S. and I. Duenyas. Optimal admission control and sequencing in a make-to-stock/make-to-order production system. *Operations Research*, 48, pp 709-720.

[6] Chan, L.M.A., D. Simchi-Levi and J. Swann. 2002. Dynamic pricing strategies for manufacturing with stochastic demand and discretionary sales. Working Paper, Massachusetts Institute of Technology.

[7] Cohen, M.A., P.R. Kleindorfer and H.L. Lee. 1988. Service constrained (s,S) inventory systems with priority demand classes and lost sales. *Management Science*, Vol. 34, pp. 482-499.

[8] de Vericourt F., F. Karaesmen and Y. Dallery. 2002. Optimal stock allocation for a capacitated supply system. *Management Science*, 48, pp 1486-1501.

[9] Duenyas, I. and M. Van Oyen. 1996. Heuristic scheduling problem with random demand. Management Science. Vol. 42, pp. 814-829.

[10] Duenyas, I., W.J. Hopp and Y. Bassok. 1997. Production quotes as bounds on interplant JIT contracts. Management Science. Vol. 43, pp. 1372-1386.

[11] Federgruen A. and Z.Katalan. 1996. The impact of setup times on the performance of multiclass service and production systems. *Operations Research*, Vol. 44, pp. 989-1001.

[12] Federguen, A. and Zipkin. 1986. An inventory model with limited production capacity and uncertain demands. I. The average-cost criterion. Mathematical Opererations Research, 11, No. 2, 193–207.

[13] Federguen, A. and Zipkin. 1986. An inventory model with limited production capacity and uncertain demands. II. The discounted-cost criterion. Mathematical Opererations Research, 11, No. 2, 208–215.

[14] Gallego, G. and G. van Ryzin. 1993. Optimal dynamic pricing of inventories with stochastic demand over finite horizons. Management Science, 30, pp. 999-1020.

[15] Gittins, J.C., *Multi-armed Bandit Allocation Indices*, Wiley, New York, 1989.

[16] Glasserman, P., and D. Yao. 1994. Monotone optimal control for permutable GSMPs. *Mathematics of Operations Research*, 19, pp 449-476.

[17] Ha, A. 1997. Inventory rationing in a make-to-stock production system with several demand classes and lost sales. *Management Science*, 43, pp 1093-1103.

[18] Ha, A. 1997. Stock-rationing policy for a make-to-stock production system with two priority classes and backordering. *Naval Research Logistics*, 44, pp 458-72.

[19] Ha, A. 2000. Stock-rationing in an M/Ek/1 make-to-stock queue. *Management Science*, 46, pp. 77-87.

[20] Kimes, S.E. 1989, A Tool for capacity-constrained service firms. Journal of Operations Management, 8, No. 3, pp. 338-363.

[21] Leibs, S. 2000. Ford heads the profits. *CFO The Magazine*, **16**, 9, pp. 33–35.

[22] Kleijn, M. and R. Dekker. 1998. An overview of inventrory systems with several demand classes. Report EI9838/A, Econometric Institute, Erasmus University Rotterdam, The Netherlands.

[23] Melchiors, P., R. Dekker and M.J. Kleijn. 2000. Inventory rationing in an (s, Q) inventory model with lost sales and two demand classes. *Journal of the Operational Research Society*, 51, pp 111-122.

[24] Nahmias, S. and W.S. Demmy. 1981. Operating characteristics of an inventory system with rationing. Management Science, 27, pp. 1236-1245.

[25] Petruzzi, N.C. and M. Dada. 1999. Pricing and the newsvendor model: a review with extensions. Operations Research, 47, pp. 183-194.

[26] Porteus, E. 1982. Conditions for characterizing the structure of optimal strategies in infnite-horizon dynamic programming. Journal of Optimization Theory and Applications, Vol. 36, pp. 419-431.

[27] Reiman, M. I. and L. M. Wein. Dynamic scheduling of a two-class queue with setups. *Operations Research*, 46, pp. 532-547.

[28] Sennott, L.I. 1999. Stochastic Dynamic Programming and the Control of Queueing Systems. Wiley.

[29] D.M. Topkis. 1968. Optimal ordering and rationing policies in a nonstationary dynamic inventory model with n demand classes. *Management Science*, 15, pp. 160-176.

[30] Varaiya, P., J. Walrand and C. Buyukkoc. 1985. Extensions of the multi-armed brandit problem. *IEEE Transactions Automatic Control*, Vol. AC-30, pp.426-439.

[31] Veatch, M.and L.M. Wein. 1996. Scheduling a make-to-stock queue: index policies and hedging points. *Operations Research*, 44, pp 634-647.

[32] Walrand J.. 1988. *An Introduction to Queueing Networks*, Prentice Hall, Englewood Cliffs, NJ.

[33] Weber R.R. and S. Stidha. 1987. Optimal control of service rates in networks of queues. *Advances in Applied Probability*, 19, pp. 202-218.